THE ASYMPTOTIC DISTRIBUTION AND ROBUSTNESS OF THE LIKELIHOOD RATIO AND SCORE TEST STATISTICS

E. A. Emberson

A Thesis Submitted for the Degree of PhD
at the
University of St Andrews

1995

L

The Asymptotic Distribution and Robustness of the Likelihood Ratio and Score Test Statistics

A thesis presented by

E. A. Emberson B.Sc. *Hons St. Andrews*

to the University of St. Andrews
in application for the degree of
Doctor of Philosophy.

October 10, 1994

Th

B702

To Mum and Dad ...

Many thanks are due to Peter Jupp, for patient supervision and much helpful advice, some of which was even taken, and to Matt for even greater patience, proof-reading and living through this with me. I would also like to thank Nick Rich and Tricia Heggie for timely and very valuable assistance. The work for this thesis was carried out with financial support from the Carnegie Trust for the Universities of Scotland.

I *Eleanor Avril Emberson* hereby certify that this thesis has been composed by myself, that it is a record of my own work, and that is has not been accepted in partial or complete fulfilment of any other degree or professional qualification.

Signed Date ...15/10/94..

I was admitted to the Faculty of Science of the University of St. Andrews under Ordinance General No 12 on *1st October 1989* and as a candidate for the degree of Ph.D. on *1st October 1990.*

Signed Date ...15/10/94...

I hereby certify that the candidate has fulfilled the conditions of the Resolution and Regulations appropriate to the degree of Ph.D.

Signed ... . Date ...18/10/94

In submitting this thesis to the University of St. Andrews I understand that I am giving permission for it to be made available for use in accordance with the regulations of the University Library for the time being in force, subject to any copyright vested in the work not being affected thereby. I also understand that the title and abstract will be published, and that a copy of the work may be made and supplied to any bona fide library or research worker.

Abstract

Cordeiro & Ferrari (1991) use the asymptotic expansion of Harris (1985) for the moment generating function of the score statistic to produce a generalization of Bartlett adjustment for application to the score statistic. It is shown here that Harris's expansion is not invariant under reparameterization and an invariant expansion is derived using a method based on the expected likelihood yoke. A necessary and sufficient condition for the existence of a generalized Bartlett adjustment for an arbitrary statistic is given in terms of its moment generating function. Generalized Bartlett adjustments to the likelihood ratio and score test statistics are derived in the case where the interest parameter is one-dimensional under the assumption of a mis-specified model, where the true distribution is not assumed to be that under the null hypothesis.

Contents

Chapter 1

Introduction

1.1 Outline

The likelihood ratio and score statistics are two of the most commonly used test statistics in parametric inference. For both statistics there are well-known results governing their asymptotic distributions, and considerable work has been done on Bartlett adjustment of the likelihood ratio statistic, the fundamental aim of which is to produce a modified statistic, asymptotically equivalent to the original, but having a distribution closer to the asymptotic distribution for small sample sizes. Until recently, no such correction procedure was known for the score statistic, but Cordeiro & Ferrari (1991) give a modified score statistic which does have the desired properties. Unfortunately, there is an error in the asymptotic expansion of the score statistic given by Harris (1985), on which their results are based. This does not, however, invalidate their method.

All of the work on producing modified likelihood ratio and score statistics has been done under the usual assumption that the true, underlying distribution of the random variables involved was that given by the null hypothesis. It is also of some interest to consider the asymptotic behaviour of the likelihood ratio and score statistics when this is not assumed to be the case.

Chapter 2 gives a brief outline of the derivation of the Bartlett adjustment to the likelihood ratio statistic, then considers the modified score test statistic given by Cordeiro & Ferrari (1991). This correction factor takes the form of a cubic in the score statistic, S_R, with the coefficients given by the asymptotic expansion of S_R. This procedure is referred to as *generalized Bartlett adjustment.* Most of the chapter is concerned with deriving a correct asymptotic expansion for the score statistic to replace that of Harris (1985).

Chapter 3 considers general conditions under which generalized Bartlett adjustment of a test statistic is possible, and, as an example, considers Bartlett adjustment of the weighted sum of test statistics for which Bartlett factors exist.

Chapter 4 moves on to consider Bartlett adjustment and generalized Bartlett adjustment of the score and likelihood ratio statistics when the null distribution is not assumed to be the true underlying distribution. Using the results of Kent (1982) and his general framework for considering mis-specified models, generalized Bartlett adjustments are derived for both the likelihood ratio and score statistics when the interest parameter in the model is one-dimensional.

1.2 Notation

1.2.1 Index Notation

If we have an m-dimensional random variable Z with components $Z_1, \ldots, Z_m$, then we write the linear combination of the Zs with coefficients $c^1, \ldots, c^m$ as

$$c^i Z_i = \sum_{i=1}^{m} c^i Z_i. \tag{1.1}$$

More generally, summation over any index within an expression which is repeated once as a subscript and once as a superscript is assumed, while any index which is not repeated is called a *free index*. For example,

$$c^{ijk} Z_i Z_j = \sum_{i=1}^{m} \sum_{j=1}^{m} c^{ijk} Z_i Z_j, \tag{1.2}$$

and k is a free index.

1.2.2 Index notation for a Partitioned Parameter

Suppose that we have an m-dimensional parameter, θ, with components $\theta^1, \ldots, \theta^m$ and that the random variable Z is a function of θ, so that the components of Z correspond to the components of θ. If θ is partitioned into a p-dimensional interest parameter, ψ, and a q-dimensional nuisance parameter, λ, with $p + q = m$, then we may sometimes wish to denote summation over the range of ψ or λ only. It is possible to do this in ordinary index notation. For example, let

$$\begin{aligned}\theta &= \left(\theta^1, \ldots, \theta^m\right) \\ &= \left(\psi^1, \ldots, \psi^p, \lambda^1, \ldots, \lambda^q\right), \end{aligned} \tag{1.3}$$

then if we wish to express

$$\sum_{i=0}^{p} \sum_{j=p+1}^{m} c^{ij} Z_i Z_j \tag{1.4}$$

in ordinary index notation we may define d^{ij} as

$$d^{ij} = \begin{cases} 0 & p+1 \leq i \leq m \\ 0 & 1 \leq j \leq p \\ c^{ij} & \text{otherwise,} \end{cases} \tag{1.5}$$

then (1.4) may be written

$$d^{ij} Z_i Z_j. \tag{1.6}$$

To avoid defining a large number of new functions and to make the range of summation explicit, a new notation is introduced here. Indices $\psi_1, \psi_2, \ldots$ and so on will be used to indicate summation over the range $1, \ldots, p$, i.e. the ψ part of θ, and similarly $\lambda_1, \lambda_2, \ldots$ and so on will indicate summation over the range $p+1, \ldots, m$, i.e. the λ part of θ. For example, (1.4) would be written as

$$c^{\psi_1 \lambda_1} Z_{\psi_1} Z_{\lambda_1}. \tag{1.7}$$

The indices ψ_i, λ_i should not be confused with the components of the interest and nuisance parameters, ψ^i and λ^i respectively.

1.2.3 [ ] Notation

We will write, for example,

$$\kappa_{i,j} \kappa_{k,l}[3] \tag{1.8}$$

to denote the sum of all three distinct terms obtained by rearranging the free indices within the given partition, i.e.

$$\kappa_{i,j}\kappa_{k,l} + \kappa_{i,k}\kappa_{j,l} + \kappa_{i,l}\kappa_{j,k}. \tag{1.9}$$

This may be used as part of a more complex expression, for example

$$\begin{aligned} c^{ij} c^{kl} \left(\kappa_{ij}\kappa_{k,l}[6]\right) &= c^{ij} c^{kl} \left(\kappa_{ij}\kappa_{k,l} + \kappa_{ik}\kappa_{j,l} \right. \\ &\quad + \kappa_{il}\kappa_{j,k} + \kappa_{jk}\kappa_{i,l} \\ &\quad \left. + \kappa_{jl}\kappa_{i,k} + \kappa_{kl}\kappa_{i,j}\right). \end{aligned} \tag{1.10}$$

1.2.4 Other Notation

Terms of the form

$$[x]_{O(n^{-1})}, \tag{1.11}$$

for instance, will denote the $O\left(n^{-1}\right)$ term of expression x when dealing with asymptotic expansions.

We will also use the convention that

$$\sum_{i=a}^{a-1} x(i) = 0, \tag{1.12}$$

where a is any integer and x any function.

1.3 Tensors

When we speak of parameterizations of a parametric model $\mathcal{M}$, we are implicitly considering alternative specifications of coordinates of a point on a manifold. Let $\phi = (\phi^1, \ldots, \phi^m)$ and $\tau = (\tau^1, \ldots, \tau^m)$ each represent the same point on an m-dimensional differentiable manifold, and let

$$\phi^i_{/a} = \frac{\partial \phi^i}{\partial \tau^a} \tag{1.13}$$

$$\tau^a_{/i} = \frac{\partial \tau^a}{\partial \phi^i}. \tag{1.14}$$

Let

$$T^{i_1 \ldots i_r}_{j_1 \ldots j_s}(\phi), \tag{1.15}$$

where each of the indices $i_1, \ldots, i_r$, $j_1, \ldots, j_s$ take values 1 to m, be a real-valued function of the m-dimensional parameter ϕ with the indices each representing components of the parameter. We may also regard this as an $(r+s)$-dimensional array. (1.15) is an (r, s) *tensor* if, under reparameterization, it satisfies

$$T^{a_1 \ldots a_r}_{b_1 \ldots b_s}(\tau) = T^{i_1 \ldots i_r}_{j_1 \ldots j_s}(\phi) \tau^{a_1}_{/i_1} \ldots \tau^{a_r}_{/i_r} \phi^{j_1}_{/b_1} \ldots \phi^{j_s}_{/b_s}. \tag{1.16}$$

A (0, 0) tensor is a *scalar*; (0,1) and (1,0) tensors are *vectors*. The concept of a tensor is very useful when we come to consider the behaviour of asymptotic expansions under reparameterization. For a fuller description of tensors, and other concepts from differential geometry as they relate to parametric statistical inference, see Barndorff-Nielsen (1986).

1.4 Regularity Conditions

In all of the subsequent work we will assume the following basic regularity conditions (similar to those in Kent (1982)).

1. The log-likelihood $l(\theta; x)$ is differentiable at least four times with respect to θ.

2. The maximum likelihood estimators are consistent - see Cox & Hinkley (1974, pp 288-292) for an outline of appropriate sufficient conditions for this.

3. The matrix functions
$$\frac{\partial l(\theta;x)}{\partial \theta}\frac{\partial l(\theta;x)}{\partial \theta^T}$$
and
$$\frac{\partial^2 l}{\partial \theta \partial \theta^T}$$
can be *smoothly averaged* at the true value of θ, θ_0, as defined by Kent (1982). A function $d(x;\theta)$ may be smoothly averaged at θ_0 if there exists a neighbourhood V of θ_0 such that:

   (a) $\int d(x;\theta)g(x)\,dx$ is well-defined and finite for $\theta \in V$, and continuous at $\theta = \theta_0$;

   (b) as $n \to \infty$
   $$\sup\left\{\left|\delta_n(\theta) - \int d(x;\theta)g(x)\,dx\right| : \theta \in V\right\} = o_p(1),$$
   where $\delta_n(\theta) = n^{-1}\sum d(X_j;\theta)$ with X_1,X_2,... a sequence of independent observations from $g(x)dx$.

   In the case of a mis-specified model, we require that these functions can be smoothly averaged at $\theta(g)$ as given by (4.2).

4. The average score function $\int \frac{\partial l(\theta;x)}{\partial \theta} g(y)\,dy$, where $g(y)$ is the true underlying density, is well-defined and finite.

5. The matrices H and K, as defined by (2.46) and (2.45) respectively under the null model and by (4.52) and (4.53) respectively under a mis-specified model, are positive definite. Kent (1982) points out that by definition these matrices are positive semi-definite.

6. Under a mis-specified model, for all $\theta \in \Theta$, the Fraser information, defined by (4.1), satisfies $-\infty < F(\theta) < \infty$, and the value $\theta(g)$ which maximizes $F(\theta)$ is unique and lies in the interior of Θ.

7. The existence of all relevant moment generating functions is also assumed.

Where the moment generating function does not exist the derivations may fairly easily be rewritten in terms of the characteristic function.

Chapter 2

Bartlett Adjustment under the Null Model

Given observations $\mathbf{x} = (x_1, \ldots, x_n)^T$ of independent, identically distributed random variables $\mathbf{X} = (X_1, \ldots, X_n)^T$ with underlying density $g(x)$ we may commonly summarize the data by fitting a parametric model $\mathcal{F} = \{f(x;\theta) : \theta \in \Theta\}$. Inference about $g(x)$ typically takes the form of a null hypothesis $H_0 : \theta \in \Theta_0$ to be tested against an alternative hypothesis $H_1 : \theta \in \Theta\backslash\Theta_0$ and two of the most common testing procedures are the likelihood ratio test and the score test.

If the log likelihood is $l(\theta; \mathbf{x}) = \log f(\mathbf{x}; \theta)$ and the maximum likelihood estimators of θ under H_0 and H_1 are $\tilde{\theta}$ and $\hat{\theta}$ respectively, then the likelihood ratio test statistic is defined as

$$w = 2[l(\hat{\theta}; \mathbf{x}) - l(\tilde{\theta}; \mathbf{x})] \tag{2.1}$$

and its large-sample asymptotic distribution under H_0 is well-known to be χ^2_d, where $d = \dim \Theta_1 - \dim \Theta_0$.

If we write the score as

$$U(\theta) = \frac{\partial l}{\partial \theta}(\theta; \mathbf{x}),$$

and its covariance matrix as

$$J(\theta) = \int U(\theta)U(\theta)^T g(x)\, dx,$$

where $g(x) \in \{f(x;\theta) : \theta \in \Theta_0\}$, then the score statistic is

$$U(\tilde{\theta})^T J(\tilde{\theta})^{-1} U(\tilde{\theta}),$$

the asymptotic distribution of which is also well-known to be χ^2_d.

Thus both the likelihood ratio statistic and the score statistic have asymptotic χ^2 distributions when we make the usual assumption that the underlying density, $g(x)$, is a member of the parametric family $\mathcal{F}$, specifically that

$$g(x) \in \{f(x;\theta) : \theta \in \Theta_0\}.$$

In this case, when we are considering the distribution of the two statistics under the assumption that the null hypothesis is true, we will say that we are working under the *null model*.

The use of these asymptotic results in practice depends on having a large value for the sample size n, though exactly how large a sample is needed to achieve a desired level of accuracy will depend on the model. In many situations, we may wish to use a likelihood ratio or score test on a "small" data set; hence we may seek more accurate distributional results for small sample sizes. In the cases considered here, this takes the form of a multiplicative correction factor for the test statistic itself, which results in an improvement in the order of the approximation by the χ^2-distribution.

2.1 Likelihood Ratio Test Statistic

2.1.1 Background

Suppose we have observations $x_1, \ldots, x_n$ of random variables $X_1, \ldots, X_n$, which we believe have an underlying distribution from the parametric family

$$\mathcal{F} = \{f(x;\theta); \theta \in \Theta\}, \tag{2.2}$$

and we wish to test the null hypothesis $H_0 : \psi = \psi_0$ against the alternative hypothesis $H_1 : \psi \neq \psi_0$, where the model parameter θ partitions as $\theta^T = \left(\psi^T, \lambda^T\right)$. Thus we have a parameter of interest, ψ, and a nuisance parameter, λ, the dimensions of which we will denote by p and q respectively, so that the dimension of θ is $p+q$. Let the maximum likelihood estimator of θ under H_0 be written as $\tilde{\theta}$ and that under H_1 be written as $\hat{\theta}$, then the likelihood ratio test statistic, w is given by

$$w = 2\left\{l(\hat{\theta};\mathbf{x}) - l(\tilde{\theta};\mathbf{x})\right\}. \tag{2.3}$$

Bartlett adjustment of the likelihood ratio statistic produces a statistic with an improved approximation to the asymptotic χ^2 distribution using a simple but very powerful approach. In outline, the idea is that, given the likelihood ratio statistic, w, which under the null hypothesis has a χ^2_p-distribution with error of order $O(n^{-1})$, we may write the mean of w as

$$\mathrm{E}(w) = \left\{1 + \frac{b}{n}\right\} p + O(n^{-\frac{3}{2}}). \tag{2.4}$$

A naive approach to producing a statistic which is distributed as χ^2_p with a greater level of accuracy might then be to consider the modified statistic w', where

$$w' = \left\{1 + \frac{b}{n}\right\}^{-1} w. \tag{2.5}$$

The remarkable fact is that this multiplicative correction procedure corrects not only the mean, but all of the cumulants of w, so that w' has a χ^2_p-distribution with error of order $O(n^{-\frac{3}{2}})$, or in some cases even smaller. This procedure, or perhaps more accurately, any procedure asymptotically equivalent to this, is known as Bartlett adjustment, and we refer to b (or sometimes to the whole multiplicative correction factor $\left\{1 + \frac{b}{n}\right\}^{-1}$) as a Bartlett factor.

The asymptotically equivalent procedures arise from the fact that it can be difficult to calculate the mean of w directly, so several different formulae have been suggested to calculate a suitable correction factor (see, for example, Lawley (1956) and Barndorff-Nielsen & Cox (1984)). For numerical studies of how Bartlett factors perform in some practical situations, see, for example, Eriksen (1985), Møller (1986) and Hollas (1991).

2.1.2 Outline of Derivation of Bartlett Adjustment

Expansion of the log-likelihood ratio statistic

It may be helpful in the understanding of Bartlett adjustment in the different situations described later to have a slightly more detailed account of the derivation of the Bartlett adjustment to the likelihood ratio statistic under the conditions described above. McCullagh (1987) considers Bartlett adjustment in the case of a simple null hypothesis (i.e. where q, the dimension of λ, is zero), but in fact the analysis can fairly easily be extended to the more general case. An outline of McCullagh's very thorough explanation is as follows.

We seek an asymptotic expansion for the likelihood ratio statistic in terms of log-likelihood derivatives and their cumulants. We write the joint cumulants of log-likelihood derivatives for a single observation as

$$\kappa_i = \mathrm{E}\left[\frac{\partial l}{\partial \theta^i}(\theta; X)\right] = 0, \tag{2.6}$$

$$\kappa_{ij} = \mathrm{E}\left[\frac{\partial^2 l}{\partial \theta^i \partial \theta^j}(\theta; X)\right], \tag{2.7}$$

$$\kappa_{ijk} = \mathrm{E}\left[\frac{\partial^3 l}{\partial \theta^i \partial \theta^j \partial \theta^k}(\theta; X)\right], \tag{2.8}$$

$$\kappa_{i,j} = \mathrm{E}\left[\frac{\partial l}{\partial \theta^i}(\theta; X)\frac{\partial l}{\partial \theta^j}(\theta; X)\right], \tag{2.9}$$

$$\kappa_{i,j,k,l} = \mathrm{E}\left[\frac{\partial l}{\partial \theta^i}(\theta; X)\frac{\partial l}{\partial \theta^j}(\theta; X)\frac{\partial l}{\partial \theta^k}(\theta; X)\frac{\partial l}{\partial \theta^l}(\theta; X)\right] - \kappa_{i,j}\kappa_{k,l}[3], \tag{2.10}$$

$$\kappa_{i,jk} = \mathrm{E}\left[\frac{\partial l}{\partial \theta^i}(\theta; X)\frac{\partial^2 l}{\partial \theta^j \partial \theta^k}(\theta; X)\right], \tag{2.11}$$

etc.

Also, we define the following random variables

$$Z_i = n^{-\frac{1}{2}}\frac{\partial l}{\partial \theta^i}(\theta; \mathbf{X}), \tag{2.12}$$

$$Z_{ij} = n^{-\frac{1}{2}}\left\{\frac{\partial^2 l}{\partial \theta^i \partial \theta^j}(\theta; \mathbf{X}) - n\kappa_{ij}\right\}, \tag{2.13}$$

$$Z_{ijk} = n^{-\frac{1}{2}}\left\{\frac{\partial^3 l}{\partial \theta^i \partial \theta^j \partial \theta^k}(\theta; \mathbf{X}) - n\kappa_{ijk}\right\}, \tag{2.14}$$

etc.

The Zs have been defined in such a way that they are all of order $O_p(1)$ and all have mean zero, and the κs have similarly been defined so that they are all O(1), so the dependence on the sample size n in the expansion will be explicit. We seek an expansion for w in terms of these Zs and κs: McCullagh shows that

$$\begin{aligned}\frac{1}{2}w = {} & \frac{1}{2}Z_r Z_s \kappa^{r,s} \\ & + n^{-\frac{1}{2}}\left\{\frac{1}{3!}\kappa_{rst}Z^r Z^s Z^t + \frac{1}{2!}Z_{rs}Z^r Z^s\right\} \\ & + n^{-1}\left\{\frac{1}{2}\left(Z_{ri}Z^i + \frac{1}{2}\kappa_{rij}Z^i Z^j\right)\kappa^{r,s}\left(Z_{si}Z^i + \frac{1}{2}\kappa_{sij}Z^i Z^j\right)\right. \\ & \left. + \frac{1}{4!}\kappa_{rstu}Z^r Z^s Z^t Z^u + \frac{1}{3!}Z_{rst}Z^r Z^s Z^t\right\} + O_p\left(n^{-\frac{3}{2}}\right),\end{aligned} \tag{2.15}$$

where $\kappa^{r,s}$ is the (r,s)-entry of the inverse of the covariance matrix of the score, $(\kappa_{r,s})$, and $Z^r = \kappa^{r,i}Z_i$.

The ϕ-parameterization

As we know that w is invariant under reparameterization, we may of course choose to work in any parameterization to simplify calculations, and, more importantly, to emphasize the underlying structure of the expansions. We can then return to our original parameterization if desired. McCullagh introduces a new parameterization by defining

$$\beta^r_{st} = \kappa^{r,i}\kappa_{i,st}, \tag{2.16}$$

$$\beta^r_{stu} = \kappa^{r,i}\kappa_{i,stu}, \tag{2.17}$$

and so on, and defining a parameter transformation from θ to ϕ in a neighbourhood of any point θ_0 by

$$\begin{aligned} \phi^r - \phi^r_0 &= \theta^r - \theta^r_0 + \beta^r_{st}(\theta^s - \theta^s_0)(\theta^t - \theta^t_0)/2! \\ &\quad + \beta^r_{stu}(\theta^s - \theta^s_0)(\theta^t - \theta^t_0)(\theta^u - \theta^u_0)/3! + \ldots. \end{aligned} \tag{2.18}$$

Then, denoting the log-likelihood derivatives in the alternative parameterization by V_r, V_{rs}, V_{rst}, etc. we find that

$$U_r = V_r, \tag{2.19}$$

$$U_{rs} = V_{rs} + \beta^i_{rs}V_i, \tag{2.20}$$

$$U_{rst} = V_{rst} + \beta^i_{rs}V_{it}[3] + \beta^i_{rst}V_i, \tag{2.21}$$

and so on. On inverting (2.19) - (2.21) we obtain

$$V_r = U_r, \tag{2.22}$$

$$V_{rs} = U_{rs} - \beta^i_{rs}U_i, \tag{2.23}$$

$$V_{rst} = U_{rst} - \beta^i_{rs}U_{it}[3] + \beta^i_{rs}\beta^j_{it}U_j[3] - \beta^i_{rst}U_i. \tag{2.24}$$

The joint cumulants in this case are denoted by ν_r, ν_{rs}, $\nu_{r,s}$ and so on, and we find that the Vs are tensors and have the property that all the covariances of V_r with the higher-order derivatives are zero, e.g.

$$\begin{aligned} \nu_{r,st} &= \mathrm{E}\left[V_r V_{st}\right] \\ &= \mathrm{E}\left[U_r\left(U_{st} - \beta^i_{st}U_i\right)\right] \end{aligned}$$

$$
\begin{aligned}
&= \kappa_{r,st} - \beta^i_{st}\kappa_{r,i} \\
&= \kappa_{r,st} - \kappa_{j,st}\kappa^{j,i}\kappa_{r,i} \\
&= 0 \qquad (2.25)
\end{aligned}
$$

and

$$
\begin{aligned}
\nu_{r,stu} &= \mathrm{E}\left[V_r V_{stu}\right] \\
&= \mathrm{E}\left[U_r\left(U_{stu} - \beta^i_{st}U_{iu}[3] + \beta^i_{st}\beta^j_{iu}U_j[3] - \beta^i_{stu}U_i\right)\right] \\
&= \kappa_{r,stu} - \beta^i_{st}\kappa_{r,iu}[3] + \beta^i_{st}\beta^j_{iu}\kappa_{r,j}[3] - \beta^i_{stu}\kappa_{r,i} \\
&= \kappa_{r,stu} - \kappa_{j,st}\kappa^{j,i}\kappa_{r,iu}[3] + \kappa_{k,st}\kappa^{k,i}\kappa_{l,iu}\kappa^{l,j}\kappa_{r,j}[3] - \kappa_{j,stu}\kappa^{i,j}\kappa_{r,i} \\
&= 0. \qquad (2.26)
\end{aligned}
$$

In general, however, it can be seen that $\nu_{rs,tu}$ and $\nu_{r,s,tu}$ are non-zero.

In our expansion for w given by equation (2.15) the expressions for the $O(1)$, $O(n^{-\frac{1}{2}})$ and $O(n^{-1})$ terms of w were each (necessarily) invariant under reparameterization, but the individual terms, e.g. $\kappa_{rst}Z^r Z^s Z^t$, were not. Working in the ϕ parameterization, where we replace all the κs by νs and redefine the Zs to be

$$Z_r = n^{-\frac{1}{2}}V_r,$$

$$Z_{rs} = n^{-\frac{1}{2}}(V_{rs} - n\nu_{rs})$$

and so on, we have the advantage that the individual terms are themselves tensors, thus each term in the sum making up the expansion is itself invariant under reparameterization.

The Effect of Bartlett Adjustment

Working in the ϕ-parameterization, McCullagh shows the mean of w to be

$$
\begin{aligned}
& p \\
+ \;& n^{-1}\left\{\frac{1}{3}\nu_{rst}\nu^{r,s,t} + \nu_{r,s,tu}\nu^{r,t}\nu^{s,u} + \frac{1}{4}\nu_{rij}\nu_{skl}\nu^{r,s}\nu^{i,j}\nu^{k,l}\right. \\
& \left. + \frac{1}{2}\nu_{rij}\nu_{skl}\nu^{r,s}\nu^{i,k}\nu^{j,l} + \nu_{rs,tu}\nu^{r,t}\nu^{s,u} + \frac{1}{4}\nu_{rstu}\nu^{r,s}\nu^{t,u}\right\} \\
+ \;& O(n^{-\frac{3}{2}}), \qquad (2.27)
\end{aligned}
$$

which is indeed of the form given by equation (2.4). Hence using the identities

$$\begin{aligned}
\nu_{rst} &= -\nu_{r,s,t}, \\
\nu_{rstu} &= -\nu_{r,s,t,u} - \nu_{r,s,tu}[6] - \nu_{rs,tu}[3],
\end{aligned}$$

and defining

$$\begin{aligned}
\rho_{13}^2 &= \nu_{i,j,k}\nu_{l,m,n}\nu^{i,j}\nu^{k,l}\nu^{m,n}, \\
\rho_{23}^2 &= \nu_{i,j,k}\nu_{l,m,n}\nu^{i,l}\nu^{j,m}\nu^{k,n}, \\
\rho_4 &= \nu_{i,j,k,l}\nu^{i,j}\nu^{k,l},
\end{aligned}$$

we have

$$\begin{aligned}
b &= \frac{1}{4p}\rho_{13}^2 + \frac{1}{6p}\rho_{23}^2 - \frac{1}{4p}\left(\nu_{r,s,t,u} - \nu_{rs,tu}[3]\right)\nu^{r,s}\nu^{t,u} \\
&\quad -\frac{1}{2p}\left(\nu_{r,s,tu} + \nu_{rs,tu}\right)\nu^{r,s}\nu^{t,u}. \qquad (2.28)
\end{aligned}$$

To show that the corrected statistic, w', given by equation (2.5), has a χ_p^2 distribution with error of order $O(n^{-\frac{3}{2}})$, we rewrite w as

$$w = W_r\nu^{r,s}W_s + O_p(n^{-\frac{3}{2}}), \qquad (2.29)$$

where

$$\begin{aligned}
W_r &= Z_r \\
&\quad + n^{-\frac{1}{2}}\left\{\frac{1}{2}Z_{rs}Z^s + \frac{1}{3!}\nu_{rst}Z^sZ^t\right\} \\
&\quad + n^{-1}\left\{\frac{1}{3!}Z_{rst}Z^sZ^t + \frac{1}{4!}\nu_{rstu}Z^sZ^tZ^u + \frac{3}{8}Z_{rs}Z^{st}Z_t \right. \\
&\qquad \left. + \frac{5}{12}Z_{rs}Z_tZ_u\nu^{stu} + \frac{1}{9}\nu_{rst}\nu_{uvw}\nu^{t,u}Z^sZ^vZ^w\right\}. \qquad (2.30)
\end{aligned}$$

The joint cumulants of the W_r up to order 4 are given by

$$\begin{aligned}
\mathrm{E}(W_r) &= \frac{1}{3!}n^{-\frac{1}{2}}\nu_{rst}\nu^{s,t} + O(n^{-\frac{3}{2}}) \qquad (2.31) \\
\mathrm{Cov}(W_r, W_s) &= \nu_{r,s} \\
&\quad + n^{-1}\left\{\frac{1}{4}\nu_{rstu}\nu^{t,u} + \nu_{rt,su}\nu^{t,u} + \nu_{r,t,su}\nu^{t,u} \right. \\
&\qquad \left. + \frac{1}{6}\nu_{r,i,j}\nu_{s,k,l}\nu^{i,k}\nu^{j,l} + \frac{2}{9}\nu_{r,s,i}\nu_{j,k,l}\nu^{i,j}\nu^{k,l}\right\} \\
&\quad + O(n^{-2}) \qquad (2.32) \\
\mathrm{Cum}(W_r, W_s, W_t) &= O(n^{-\frac{3}{2}}) \qquad (2.33) \\
\mathrm{Cum}(W_r, W_s, W_t, W_u) &= O(n^{-2}). \qquad (2.34)
\end{aligned}$$

As the higher order joint cumulants are of order $O(n^{-\frac{3}{2}})$ or smaller, the vector $\mathbf{W} = (W_1, \ldots, W_p)^T$ has, with error of order $O(n^{-\frac{3}{2}})$, a multivariate normal distribution with mean vector and covariance matrix given by equations (2.31) and (2.32) respectively. Thus to the same order of approximation w has a scaled, non-central χ_p^2 distribution, the rth cumulant of which is given by

$$2^{r-1}(r-1)!\,p\left\{1+\frac{b}{n}\right\}^r + O(n^{-\frac{3}{2}}). \tag{2.35}$$

Hence $w' = (1+\frac{b}{n})^{-1} w$ has a χ_p^2 distribution with error of order $O(n^{-\frac{3}{2}})$.

McCullagh points out that, although the formal derivation is quite general, the effect of Bartlett adjustment in the case of discrete data is not quite so simple as the above might suggest. For a discussion of the performance of Bartlett adjustment of the likelihood ratio statistic in the discrete case, see Frydenberg & Jensen (1989).

2.2 Score Test Statistic

2.2.1 Introduction

Consider now the score statistic for n independent identically distributed observations, $\mathbf{x} = (x_1, \ldots, x_n)$,

$$S_R = \tilde{Z}_i \tilde{\kappa}^{i,j} \tilde{Z}_j \tag{2.36}$$

where, as in (2.12) and (2.9),

$$Z_i = n^{-\frac{1}{2}} \frac{\partial l}{\partial \theta^i}(\theta; \mathbf{x}),$$

$$\kappa_{i,j} = \mathrm{E}\left[\frac{\partial l}{\partial \theta^i}(\theta; X)\frac{\partial l}{\partial \theta^j,}(\theta; X)\right],$$

i.e. $\kappa_{i,j}$ is the (i,j) element of the covariance matrix, $\kappa^{i,j}$ is the (i,j) element of the inverse of the covariance matrix of the score for a single observation, and the addition of a tilde indicates that functions are evaluated at the point $\tilde{\theta}$, the maximum likelihood estimate under H_0. As in the case of the likelihood ratio statistic, the score statistic is asymptotically distributed as a χ_p^2 for a test of $H_0 : \psi = \psi_0$ against $H_1 : \psi \neq \psi_0$, where $\theta^T = \left(\psi^T, \lambda^T\right)$, with ψ a p-dimensional interest parameter, λ a q-dimensional nuisance parameter, and $p + q = m$. As has been shown previously, (see section 2.1.2) we can modify the likelihood ratio statistic, w, to produce a modified test statistic, $w' = w\left\{1 + \frac{b}{n}\right\}^{-1}$, the distribution of which is χ_p^2 with error of order $O\left(n^{-\frac{3}{2}}\right)$. However, there is no equivalent multiplicative correction factor for the score statistic (see Harris (1985)).

2.2.2 A Generalized Bartlett Adjustment

Although we cannot find a simple multiplicative correction factor of the above form, it is possible to find an improved test statistic by a different method. Cordeiro and Ferrari (1991) use an expansion derived by Harris (1985) to obtain a modified score test statistic which is distributed as a χ_p^2 with error of order $O(n^{-\frac{3}{2}})$. This modified statistic is of the form

$$S'_R = S_R\left\{1 - \left(\gamma + \beta S_R + \alpha S_R^2\right)\right\},$$

where the coefficients α, β and γ are functions of the joint cumulants of the log-likelihood derivatives up to order 4. The idea comes from the asymptotic expansion, to order $O\left(n^{-1}\right)$, for the moment generating function of the null distribution of S_R. Harris gives this as

$$M(t) = (1-2t)^{-\frac{p}{2}}\left\{1 + (24n)^{-1}\left(A_1 d + A_2 d^2 + A_3 d^3\right)\right\} + o\left(n^{-1}\right), \qquad (2.37)$$

where

$$d = \frac{2t}{1-2t}$$

and A_1, A_2 and A_3 are functions of the joint cumulants of log-likelihood derivatives. We can rewrite (2.37) as

$$\begin{aligned} M(t) &= (1-2t)^{-\frac{p}{2}} \\ &\quad + (24n)^{-1}\left\{A_3(1-2t)^{-\frac{p+6}{2}} + (A_2 - 3A_3)(1-2t)^{-\frac{p+4}{2}}\right. \\ &\qquad + (3A_3 - 2A_2 + A_1)(1-2t)^{-\frac{p+2}{2}} \\ &\qquad \left. + (A_2 - A_1 - A_3)(1-2t)^{-\frac{p}{2}}\right\} \\ &\quad + o\left(n^{-1}\right), \end{aligned} \qquad (2.38)$$

from which it follows that

$$\begin{aligned} f_{S_R}(x) &= g_p(x) + (24n)^{-1}\left\{A_3 g_{p+6}(x) + (A_2 - 3A_3)\, g_{p+4}(x)\right. \\ &\quad \left. + (3A_3 - 2A_2 + A_1)\, g_{p+2}(x) + (A_2 - A_1 - A_3)\, g_p(x)\right\} \\ &\quad + o\left(n^{-1}\right), \end{aligned} \qquad (2.39)$$

where $f_{S_R}(x)$ is the probability density function of S_R and $g_p(x)$ is the probability density function of a χ_p^2 random variable. Now from the recurrence relation $g_{p+2}(x) = xp^{-1}g_p(x)$ we have

$$f_{S_R}(x) = g_p(x)\left(1 + B_0 + B_1x + B_2x^2 + B_3x^3\right) + o(n^{-1}), \tag{2.40}$$

where

$$\begin{aligned}
B_0 &= \frac{A_2 - A_1 - A_3}{24n}, \\
B_1 &= \frac{3A_3 - 2A_2 + A_1}{24pn}, \\
B_2 &= \frac{A_2 - 3A_3}{24p\,(p+2)\,n}, \\
B_3 &= \frac{A_3}{24p\,(p+2)\,(p+4)\,n}.
\end{aligned}$$

Cordeiro and Ferrari (1991) use the above results to obtain a modified score statistic: their result may be expressed as follows.

Proposition 2.1 *If a statistic S_R has a moment generating function of the form (2.38) then the statistic S_R' given by*

$$S_R' = S_R\left\{1 - (\gamma + \beta S_R + \alpha S_R^2)\right\}, \tag{2.41}$$

where

$$\alpha = \frac{A_3}{12p(p+2)(p+4)n}, \tag{2.42}$$

$$\beta = \frac{A_2 - 2A_3}{12p(p+2)n}, \tag{2.43}$$

$$\gamma = \frac{A_1 - A_2 + A_3}{12pn}, \tag{2.44}$$

has a χ_p^2 distribution with error of order $o(n^{-1})$.

Proof : see Cordeiro & Ferrari (1991).

2.2.3 Behaviour of Expansions under Reparameterization

We must now consider the expansion (2.37) given by Harris in more detail. In the interest of consistency, we follow the notation of McCullagh (1987) as used previously instead of the notation used by Harris in his paper: thus the joint cumulants of the log-likelihood derivatives are written as $\kappa_{i,j}$, κ_{ij}, and so on, as defined in section 2.1.2. Because of the partition of the m-dimensional parameter θ into a p-dimensional interest parameter, ψ, and a q-dimensional nuisance parameter, λ, it is also useful in what follows to express the covariance of the score and the expected second derivative of the log-likelihood as partitioned matrices, i.e.

$$\begin{aligned} K &= (\kappa_{i,j})_{m\times m} \\ &= \begin{bmatrix} K_{\psi,\psi} & K_{\psi,\lambda} \\ K_{\lambda,\psi} & K_{\lambda,\lambda} \end{bmatrix} \qquad (2.45) \\ H &= (\kappa_{ij})_{m\times m} \\ &= \begin{bmatrix} H_{\psi\psi} & H_{\psi\lambda} \\ H_{\lambda\psi} & H_{\lambda\lambda} \end{bmatrix}. \qquad (2.46) \end{aligned}$$

A fundamental property of the score statistic, and indeed of any sensible test statistic, is invariance under reparameterization. The expressions given by Harris for A_1, A_2 and A_3 are (correcting a misprint noted by Cordeiro and Ferrari (1991))

$$\begin{aligned} A_1 &= 3(\kappa_{ijk} + 2\kappa_{i,jk})(\kappa_{rst} + 2\kappa_{rs,t})a^{ij}a^{st}m^{kr} \\ &\quad -6(\kappa_{ijk} + \kappa_{i,jk})\kappa_{r,s,t}a^{ij}a^{kr}m^{st} \\ &\quad +6(\kappa_{i,jk} - \kappa_{i,j,k})(\kappa_{rst} + 2\kappa_{rs,t})a^{js}a^{kt}m^{ir} \\ &\quad -6(\kappa_{i,j,k,l} + \kappa_{i,j,kl})a^{kl}m^{ij}, \qquad (2.47) \\ A_2 &= -3\kappa_{i,j,k}\kappa_{r,s,t}a^{kr}m^{ij}m^{st} \\ &\quad +6(\kappa_{ijk} + 2\kappa_{i,jk})\kappa_{r,s,t}a^{ij}m^{kr}m^{st} \\ &\quad -6\kappa_{i,j,k}\kappa_{r,s,t}a^{kt}m^{ir}m^{js} \\ &\quad +3\kappa_{i,j,k,l}m^{ij}m^{kl}, \qquad (2.48) \\ A_3 &= 3\kappa_{i,j,k}\kappa_{r,s,t}m^{ij}m^{kr}m^{st} \\ &\quad +2\kappa_{i,j,k}\kappa_{r,s,t}m^{ir}m^{js}m^{kt}, \qquad (2.49) \end{aligned}$$

where m^{ij} and a^{ij} denote the (i,j) elements of the matrices M and A respectively, defined by

$$\begin{aligned} A &= \begin{bmatrix} 0 & 0 \\ 0 & K_{\lambda,\lambda}^{-1} \end{bmatrix}, \qquad (2.50) \\ M &= K^{-1} - A. \qquad (2.51) \end{aligned}$$

As the score statistic S_R is invariant under reparameterization, it follows that A_1,

A_2 and A_3 should also be invariant, but further consideration of equations (2.47) and (2.48) show that this is not the case as we shall see.

To investigate the behaviour of A_1, A_2 and A_3 under a change of parameters, consider a new parameterization $\omega = (\phi, \mu)$ where, in general, $\phi = \phi(\psi)$ and $\mu = \mu(\psi, \lambda)$. Denoting log-likelihood cumulants in this parameterization by $\overline{\kappa}_{i,j}$, $\overline{\kappa}_{ijk}$, etc. we have

$$\begin{aligned}
\overline{\kappa}_{i,j} &= \mathrm{E}\left[\frac{\partial l}{\partial \omega^i}(\omega; X)\frac{\partial l}{\partial \omega^j}(\omega; X)\right] \\
&= \mathrm{E}\left[\frac{\partial \theta^r}{\partial \omega^i}\frac{\partial l}{\partial \theta^r}(\theta; X)\frac{\partial \theta^s}{\partial \omega^j}\frac{\partial l}{\partial \theta^s}(\theta; X)\right] \\
&= \theta^r_{/i}\theta^s_{/j}\kappa_{r,s},
\end{aligned} \tag{2.52}$$

where

$$\theta^r_{/i} = \frac{\partial \theta^r}{\partial \omega^i}.$$

Similarly

$$\overline{\kappa}_{i,j,k} = \theta^r_{/i}\theta^s_{/j}\theta^t_{/k}\kappa_{r,s,t},$$

$$\overline{\kappa}_{i,j,k,l} = \theta^r_{/i}\theta^s_{/j}\theta^t_{/k}\theta^u_{/l}\kappa_{r,s,t,u},$$

and, since $\kappa_{ij} = -\kappa_{i,j}$,

$$\overline{\kappa}_{ij} = \theta^r_{/i}\theta^s_{/j}\kappa_{rs},$$

but

$$\begin{aligned}
\overline{\kappa}_{i,jk} &= \mathrm{E}\left[\frac{\partial l}{\partial \omega^i}(\omega; X)\frac{\partial^2 l}{\partial \omega^j \partial \omega^k}(\omega; X)\right] \\
&= \mathrm{E}\left[\frac{\partial \theta^r}{\partial \omega^i}\frac{\partial l}{\partial \theta^r}(\theta; X)\frac{\partial}{\partial \omega^k}\left\{\frac{\partial \theta^s}{\partial \omega^j}\frac{\partial l}{\partial \theta^s}(\theta; X)\right\}\right] \\
&= \mathrm{E}\left[\frac{\partial \theta^r}{\partial \omega^i}\frac{\partial l}{\partial \theta^r}(\theta; X)\left\{\frac{\partial \theta^s}{\partial \omega^j}\frac{\partial \theta^t}{\partial \omega^k}\frac{\partial^2 l}{\partial \theta^s \partial \theta^t}(\theta; X) + \frac{\partial^2 \theta^s}{\partial \omega^j \partial \omega^k}\frac{\partial l}{\partial \theta^s}(\theta; X)\right\}\right] \\
&= \theta^r_{/i}\theta^s_{/j}\theta^t_{/k}\kappa_{r,st} + \theta^r_{/i}\theta^s_{/jk}\kappa_{r,s},
\end{aligned} \tag{2.53}$$

where

$$\theta^r_{/ij} = \frac{\partial^2 \theta^r}{\partial \omega^i \partial \omega^j}.$$

We also find that

$$\overline{\kappa}_{ijk} = \mathrm{E}\left[\frac{\partial^3 l}{\partial \omega^i \partial \omega^j \partial \omega^k}(\omega; X)\right]$$

$$
\begin{aligned}
&= \mathrm{E}\left[\frac{\partial^2}{\partial\omega^j\partial\omega^k}\left\{\frac{\partial\theta^r}{\partial\omega^i}\frac{\partial l}{\partial\theta^r}(\theta;X)\right\}\right] \\
&= \mathrm{E}\left[\frac{\partial}{\partial\omega^k}\left\{\frac{\partial\theta^r}{\partial\omega^i}\frac{\partial\theta^s}{\partial\omega^j}\frac{\partial^2 l}{\partial\theta^r\partial\theta^s}(\theta;X)+\frac{\partial^2\theta^r}{\partial\omega^i\partial\omega^j}\frac{\partial l}{\partial\theta^r}(\theta;X)\right\}\right] \\
&= \mathrm{E}\left[\frac{\partial\theta^r}{\partial\omega^i}\frac{\partial\theta^s}{\partial\omega^j}\frac{\partial\theta^t}{\partial\omega^k}\frac{\partial^3 l}{\partial\theta^r\partial\theta^s\partial\theta^t}(\theta;X)+\frac{\partial\theta^r}{\partial\omega^i}\frac{\partial^2\theta^s}{\partial\omega^j\partial\omega^k}\frac{\partial^2 l}{\partial\theta^r\partial\theta^s}(\theta;X)[3]\right. \\
&\qquad \left.+\frac{\partial^3\theta^r}{\partial\omega^i\partial\omega^j\partial\omega^k}\frac{\partial l}{\partial\theta^r}(\theta;X)\right] \\
&= \theta^r_{/i}\theta^s_{/j}\theta^t_{/k}\kappa_{rst}+\theta^r_{/i}\theta^s_{/jk}\kappa_{rs}[3]
\end{aligned}
$$

and

$$
\begin{aligned}
\overline{\kappa}_{i,j,kl} &= \mathrm{E}\left[\frac{\partial l}{\partial\omega^i}(\omega;X)\frac{\partial l}{\partial\omega^j}(\omega;X)\frac{\partial^2 l}{\partial\omega^k\partial\omega^l}(\omega;X)\right] \\
&= \mathrm{E}\left[\frac{\partial\theta^r}{\partial\omega^i}\frac{\partial l}{\partial\theta^r}(\theta;X)\frac{\partial\theta^s}{\partial\omega^j}\frac{\partial l}{\partial\theta^s}(\theta;X)\frac{\partial}{\partial\omega^l}\left\{\frac{\partial\theta^t}{\partial\omega^k}\frac{\partial l}{\partial\theta^t}(\theta;X)\right\}\right] \\
&= \mathrm{E}\left[\frac{\partial\theta^r}{\partial\omega^i}\frac{\partial l}{\partial\theta^r}(\theta;X)\frac{\partial\theta^s}{\partial\omega^j}\frac{\partial l}{\partial\theta^s}(\theta;X)\right. \\
&\qquad \left.\left\{\frac{\partial\theta^t}{\partial\omega^k}\frac{\partial\theta^u}{\partial\omega^l}\frac{\partial^2 l}{\partial\theta^t\partial\theta^u}(\theta;X)+\frac{\partial^2\theta^t}{\partial\omega^k\partial\omega^l}\frac{\partial l}{\partial\theta^t}(\theta;X)\right\}\right] \\
&= \theta^r_{/i}\theta^s_{/j}\theta^t_{/k}\theta^u_{/l}\kappa_{r,s,tu}+\theta^r_{/i}\theta^s_{/j}\theta^t_{/kl}\kappa_{r,s,t}.
\end{aligned}
$$

Note that we may consider $\theta^r_{/i}$ as the (r,i) element of an $m\times m$ matrix, and denote the (i,s) element of the inverse matrix by $\omega^i_{/s}$, where

$$\theta^r_{/i}\omega^i_{/s}=\delta^i_k,$$

and

$$\delta^i_k=\begin{cases}1 & \text{if } i=k\\ 0 & \text{otherwise.}\end{cases}$$

Note that m^{ij} and a^{ij} also transform tensorially i.e.

$$
\begin{aligned}
\overline{m}^{ij} &= \omega^i_{/r}\omega^j_{/s}m^{rs} \\
\overline{a}^{ij} &= \omega^i_{/r}\omega^j_{/s}a^{rs}.
\end{aligned}
$$

From these transformation rules it is easy to see that A_3 is invariant under reparameterization. However, when we consider the effect of a parameter transformation on A_2, as given by (2.48) we find

$$\overline{A}_2 = -3\overline{\kappa}_{i,j,k}\overline{\kappa}_{r,s,t}\overline{a}^{kr}\overline{m}^{ij}\overline{m}^{st}$$

$$
\begin{aligned}
& +6(\overline{\kappa}_{ijk} + 2\overline{\kappa}_{i,jk})\overline{\kappa}_{r,s,t}\overline{a}^{ij}\overline{m}^{kr}\overline{m}^{st} \\
& -6\overline{\kappa}_{i,j,k}\overline{\kappa}_{r,s,t}\overline{a}^{kt}\overline{m}^{ir}\overline{m}^{js} \\
& +3\overline{\kappa}_{i,j,k,l}\overline{m}^{ij}\overline{m}^{kl} \\
= \; & -3\kappa_{i,j,k}\kappa_{r,s,t}a^{kr}m^{ij}m^{st} \\
& -6\kappa_{i,j,k}\kappa_{r,s,t}a^{kt}m^{ir}m^{js} \\
& +3\kappa_{i,j,k,l}m^{ij}m^{kl} \\
& +6(\theta^u_{/i}\theta^v_{/j}\theta^w_{/k}\kappa_{uvw} + \theta^u_{/i}\theta^v_{/jk}\kappa_{uv}[3] + 2\theta^u_i\theta^v_j\theta^w_k\kappa_{u,vw} + 2\theta^u_{/i}\theta^v_{/jk}\kappa_{u,v}) \\
& \theta^a_{/r}\theta^b_{/s}\theta^c_{/t}\kappa_{a,b,c}\omega^i_{/d}\omega^j_{/e}a^{de}\omega^k_{/f}\omega^r_{/g}m^{fg}\omega^s_{/x}\omega^t_{/y}m^{xy} \\
= \; & A_2 \\
& +6\left\{\omega^i_{/d}\omega^j_{/e}\omega^k_{/f}(\theta^u_{/i}\theta^v_{/jk}\kappa_{uv}[3]) + 2\omega^j_{/e}\omega^k_{/f}\theta^v_{/jk}\kappa_{d,v}\right\}\kappa_{a,b,c}a^{de}m^{fa}m^{bc} \\
\neq \; & A_2 \text{ in general.}
\end{aligned}
\tag{2.54}
$$

Similarly we can show that for A_1 as given by (2.47),

$$\overline{A}_1 \neq A_1 \text{ in general.} \tag{2.55}$$

Thus the expressions for A_1 and A_2 given by Harris are incorrect. Unfortunately, owing to restrictions on space, Harris's paper does not include any detail of the calculations involved in the derivation of the expressions for A_1, A_2 and A_3, so it was necessary to recalculate the expansion in order to find the error. The rest of this chapter gives details of the recalculation.

2.2.4 Basic Taylor Expansions

In order to calculate $M(t)$, it is necessary first to obtain an expansion for S_R. As before (see section 2.1.2 on the likelihood ratio statistic), for $\mathbf{X} = (X_1, \dots, X_n)$, consider the random variables Z_i, Z_{ij}, Z_{ijk} as defined by (2.12), (2.13), (2.14) respectively. We seek an expansion for S_R in terms of the Zs and the joint cumulants of the log-likelihood derivatives.

Looking firstly at $\tilde{Z}_i$ and expanding in a Taylor series about the true value θ we find that

$$
\begin{aligned}
\tilde{Z}_i &= n^{-\frac{1}{2}}\frac{\partial l}{\partial \theta^i}(\tilde{\theta}; X) \\
&= n^{-\frac{1}{2}}\frac{\partial l}{\partial \theta^i}(\theta; X) + n^{-\frac{1}{2}}(\tilde{\theta} - \theta)^j \frac{\partial^2 l}{\partial \theta^i \partial \theta^j}(\theta; X)
\end{aligned}
$$

$$
\begin{aligned}
& +\frac{1}{2!}n^{-\frac{1}{2}}(\tilde{\theta}-\theta)^{j}(\tilde{\theta}-\theta)^{k}\frac{\partial^3 l}{\partial\theta^i\partial\theta^j\partial\theta^k}(\theta;X)+\ldots \\
= \;& Z_i+\left(Z_{ij}+n^{\frac{1}{2}}\kappa_{ij}\right)n^{-\frac{1}{2}}\tilde{\delta}^j \\
& +\frac{1}{2}\left(Z_{ijk}+n^{\frac{1}{2}}\kappa_{ijk}\right)n^{-1}\tilde{\delta}^j\tilde{\delta}^k \\
& +\frac{1}{6}\left(Z_{ijkl}+n^{\frac{1}{2}}\kappa_{ijkl}\right)n^{-\frac{3}{2}}\tilde{\delta}^j\tilde{\delta}^k\tilde{\delta}^l+\ldots \\
= \;& Z_i+\kappa_{ij}\tilde{\delta}^j+n^{-\frac{1}{2}}\left(Z_{ij}\tilde{\delta}^j+\frac{1}{2}\kappa_{ijk}\tilde{\delta}^j\tilde{\delta}^k\right) \\
& +n^{-1}\left(\frac{1}{2}Z_{ijk}\tilde{\delta}^j\tilde{\delta}^k+\frac{1}{6}\kappa_{ijkl}\tilde{\delta}^j\tilde{\delta}^k\tilde{\delta}^l\right)+O(n^{-\frac{3}{2}}), \qquad (2.56)
\end{aligned}
$$

where

$$
\tilde{\delta}^i = n^{\frac{1}{2}}(\tilde{\theta}-\theta)^{i},
$$

so that $\tilde{\delta}^i$ is of order $O_p(1)$. Similarly we can Taylor expand $\tilde{\kappa}_{i,j}$ about θ to obtain

$$
\tilde{\kappa}_{i,j}=\kappa_{i,j}+\left(\frac{\partial}{\partial\theta^k}\kappa_{i,j}\right)n^{-\frac{1}{2}}\tilde{\delta}^k+\frac{1}{2!}\left(\frac{\partial^2}{\partial\theta^k\partial\theta^l}\kappa_{i,j}\right)n^{-1}\tilde{\delta}^k\tilde{\delta}^l+O(n^{-\frac{3}{2}}). \qquad (2.57)
$$

Now using the result that, for $m \times m$ matrices A and H, with A invertible

$$
\left(A+n^{-\frac{1}{2}}H\right)^{-1}=A^{-1}-n^{-\frac{1}{2}}A^{-1}HA^{-1}+n^{-1}A^{-1}HA^{-1}HA^{-1}+\ldots, \qquad (2.58)
$$

we find

$$
\begin{aligned}
\tilde{\kappa}^{i,j} \;=\;& \kappa^{i,j} \\
& -n^{-\frac{1}{2}}\kappa^{i,r}\left[\left(\frac{\partial}{\partial\theta^a}\kappa_{r,s}\right)\tilde{\delta}^a+n^{-\frac{1}{2}}\frac{1}{2}\left(\frac{\partial^2}{\partial\theta^a\partial\theta^b}\kappa_{r,s}\right)\tilde{\delta}^a\tilde{\delta}^b\right]\kappa^{s,j} \\
& +n^{-1}\kappa^{i,r}\left(\frac{\partial}{\partial\theta^a}\kappa_{r,s}\right)\tilde{\delta}^a\kappa^{s,t}\left(\frac{\partial}{\partial\theta^b}\kappa_{t,u}\right)\tilde{\delta}^b\kappa^{u,j}+O\left(n^{-\frac{3}{2}}\right). \qquad (2.59)
\end{aligned}
$$

2.2.5 Differentiation of Tensors

An ordinary derivative of a tensor in an arbitrary coordinate system is not itself a tensor in general. Consider, for example,

$$
\begin{aligned}
\frac{\partial}{\partial\theta^k}\kappa_{i,j} \;=\;& \frac{\partial}{\partial\theta^k}\int\frac{\partial l}{\partial\theta^i}(\theta;x)\frac{\partial l}{\partial\theta^j}(\theta;x)f(x;\theta)\,dx \\
=\;& \int\frac{\partial^2 l}{\partial\theta^i\partial\theta^k}(\theta;x)\frac{\partial l}{\partial\theta^j}(\theta;x)f(x;\theta)\,dx \\
& +\int\frac{\partial l}{\partial\theta^i}(\theta;x)\frac{\partial^2 l}{\partial\theta^j\partial\theta^k}(\theta;x)f(x;\theta)\,dx \\
& +\int\frac{\partial l}{\partial\theta^i}(\theta;x)\frac{\partial l}{\partial\theta^j}(\theta;x)\frac{\partial l}{\partial\theta^k}(\theta;x)f(x;\theta)\,dx \\
=\;& \kappa_{j,ik}+\kappa_{i,jk}+\kappa_{i,j,k}, \qquad (2.60)
\end{aligned}
$$

which clearly does not transform tensorially. In order to understand properly what is meant by differentiation in this context, we must consider *covariant differentiation* (for a full account of this in a statistical context see e.g. Murray & Rice (1993)).

Firstly we define a *connection* ∇ as a mapping

$$\begin{aligned}\nabla : \mathcal{X}(M) \times \mathcal{X}(M) &\to \mathcal{X}(M) \\ (X, Y) &\to \nabla_X Y\end{aligned}$$

where $\mathcal{X}(M)$ is the set of smooth vector fields on the differentiable manifold M, and ∇ satisfies the following:

$$\nabla_{fX+gY} Z = f \nabla_X Z + g \nabla_Y Z \tag{2.61}$$

$$\nabla_X (fY + gZ) = f \nabla_X Y + (Xf)Y + g \nabla_X Z + (Xg)Z \tag{2.62}$$

for all $f, g \in C^\infty(M)$, and $X, Y, Z \in \mathcal{X}(M)$. Then we note that $(\nabla_X Y)_p$, the value of $\nabla_X Y$ at p, depends on X only through X_p, i.e.

$$(\nabla_X Y)_p = \nabla_{X_p} Y \tag{2.63}$$

and this quantity is referred to as the covariant derivative of Y along the tangent vector X_p to M at p. If ω is a cotangent vector (1-form), then the covariant derivatives of $Y \otimes Z$ and ω are given by

$$\nabla_{X_p} (Y \otimes Z) = \left(\nabla_{X_p} Y\right) \otimes Z + Y \otimes \left(\nabla_{X_p} Z\right) \tag{2.64}$$

and

$$\left(\nabla_{X_p} \omega\right)(Y) = \omega(Y) - \omega\left(\nabla_{X_p} Y\right). \tag{2.65}$$

Covariant differentiation of higher order tensors is defined in a similar way.

Given local coordinates $(\phi^1, \ldots, \phi^m)$ around p, a connection ∇ may be completely specified by the *(upper) Christoffel symbols*, Γ^k_{ij}, defined as

$$\nabla_{\frac{\partial}{\partial \phi^j}} \frac{\partial}{\partial \phi^i} = \Gamma^k_{ij} \frac{\partial}{\partial \phi^k}. \tag{2.66}$$

In coordinate terms, the covariant derivative of a tensor $A^{i_1 \ldots i_r}_{j_1 \ldots j_s}$ with respect to x^k may be written as

$$\begin{aligned}\nabla_{\frac{\partial}{\partial x^k}} A^{i_1 \ldots i_r}_{j_1 \ldots j_s} &= \frac{\partial}{\partial x^k} A^{i_1 \ldots i_r}_{j_1 \ldots j_s} + \sum_{\alpha=1}^{r} \Gamma^{i_\alpha}_{lk} A^{i_1 \ldots i_{\alpha-1} l i_{\alpha+1} \ldots i_r}_{j_1 \ldots j_s} \\ &\quad - \sum_{\beta=1}^{s} \Gamma^l_{j_\beta k} A^{i_1 \ldots i_r}_{j_1 \ldots j_{\beta-1} l j_{\beta+1} \ldots j_s}.\end{aligned} \tag{2.67}$$

A *flat connection* is one for which, in some coordinate system, $\Gamma_{ij}^k = 0$ everywhere. Covariant differentiation with respect to a flat connection reduces to ordinary differentiation, as can be seen from (2.67).

In a vector space, there is a unique flat connection given by $\Gamma_{ij}^k = 0$ in the coordinate system given by a basis. Thus differentiation in any such coordinate system will yield Taylor expansions which are parameterization-independent. Where we do not have a vector space, differentiation must be handled carefully to produce invariant asymptotic expansions. It is not enough to use covariant differentiation, as in order to use Taylor expansion, the various terms must be ordinary covariant derivatives in some coordinate system and these are not parameterization-invariant. The solution is to use a coordinate system based on the *expected likelihood yoke* (see Barndorff-Nielsen (1987) and Barndorff-Nielsen, Jupp & Kendall (1994)).

This is, in fact, the ϕ-parameterization explained previously (see page 11). Working in this parameterization (where, for distinction, we replace the κs by νs), the derivatives of the tensors in (2.57) become

$$\begin{aligned} \frac{\partial}{\partial \phi^k}\nu_{i,j} &= \nu_{i,j,k} + \nu_{i,jk} + \nu_{ik,j} \\ &= \nu_{i,j,k}, \end{aligned} \tag{2.68}$$

since $\nu_{i,jk} = \nu_{ik,j} = 0$. Also,

$$\begin{aligned} \frac{\partial^2}{\partial \phi^k \partial \phi^l}\nu_{i,j} &= \frac{\partial}{\partial \phi^l}\nu_{i,j,k} \\ &= \nu_{i,j,kl} + \nu_{i,k,jl} + \nu_{j,k,il} + \nu_{i,j,k,l}. \end{aligned} \tag{2.69}$$

The use of this parameterization in fact simplifies many asymptotic calculations (see for instance McCullagh's (1987) derivation of the Bartlett adjustment for the likelihood ratio test statistic, outlined in section 2.1.2).

2.2.6 Expansion of S_R in the ϕ-Parameterization

When using this parameterization, we will write $\phi = (\tau, \xi)$, where τ is the p-dimensional interest parameter and ξ is the q-dimensional nuisance parameter, and distinguish the new versions of the matrices A, K and H by using ϕ, τ and ξ as sub- and superscripts. The equation (2.59) may now be written as

$$\tilde{\nu}^{i,j} = \nu^{i,j}$$

$$
\begin{aligned}
&-n^{-\frac{1}{2}}\nu^{i,k}\nu_{k,l,a}\tilde{\delta}^a\nu^{l,j} \\
&+n^{-1}\left[\frac{1}{2}\nu^{i,k}(\nu_{k,l,ab}+\nu_{k,a,lb}+\nu_{l,a,kb}+\nu_{k,l,a,b})\tilde{\delta}^a\tilde{\delta}^b\nu^{l,j}\right. \\
&\qquad\left.+\nu^{i,k}\nu_{k,l,a}\tilde{\delta}^a\nu^{l,m}\nu_{m,n,b}\tilde{\delta}^b\nu^{n,j}\right] \\
&+O(n^{-\frac{3}{2}}).
\end{aligned} \tag{2.70}
$$

Since ϕ is the true value of the parameter under the null hypothesis, we have

$$
\begin{aligned}
\tilde{\delta}^i &= (\tilde{\phi}-\phi)^i \\
&= 0, \text{ for } i=1,\ldots,p,
\end{aligned}
$$

i.e. the τ part of $\tilde{\delta}$ is zero. Also

$$
\begin{aligned}
\tilde{Z}_i &= \frac{\partial l}{\partial \phi^i}(\tilde{\phi};X) \\
&= 0, \text{ for } i=p+1,\ldots,m,
\end{aligned} \tag{2.71}
$$

i.e. the ξ part of $\tilde{Z}$ is zero. Substituting into (2.56) we have (now using index notation for partitioned parameter, see section 1.2.2)

$$
\begin{aligned}
-Z_{\xi_1} &= \nu_{\xi_1\xi_2}\tilde{\delta}^{\xi_2} \\
&\quad+n^{-\frac{1}{2}}\left(Z_{\xi_1\xi_2}\tilde{\delta}^{\xi_2}+\frac{1}{2}\nu_{\xi_1\xi_2\xi_3}\tilde{\delta}^{\xi_2}\tilde{\delta}^{\xi_3}\right) \\
&\quad+n^{-1}\left(\frac{1}{2}Z_{\xi_1\xi_2\xi_3}\tilde{\delta}^{\xi_2}\tilde{\delta}^{\xi_3}+\frac{1}{6}\nu_{\xi_1\xi_2\xi_3\xi_4}\tilde{\delta}^{\xi_2}\tilde{\delta}^{\xi_3}\tilde{\delta}^{\xi_4}\right) \\
&\quad+O(n^{-\frac{3}{2}}).
\end{aligned} \tag{2.72}
$$

In order to invert equation (2.72), we use the following result.

Lemma 2.1 *If we have an invertible function*

$$
T(v)=T_1(v)+\frac{1}{2}T_2(v,v)+\frac{1}{6}T_3(v,v,v)+\ldots, \tag{2.73}
$$

where $v\in E$, for some vector space E, and if we denote the inverse function by $R(v)$, where

$$
R(v)=R_1(v)+\frac{1}{2}R_2(v,v)+\frac{1}{6}R_3(v,v,v)+\ldots, \tag{2.74}
$$

then

$$
R_1(v) = T_1^{-1}(v) \tag{2.75}
$$

$$
R_2(v,v) = -T_1^{-1}(T_2(T_1^{-1}(v),T_1^{-1}(v))) \tag{2.76}
$$

$$
R_3(v,v,v) = -T_1^{-1}(T_3(T_1^{-1}(v),T_1^{-1}(v),T_1^{-1}(v))). \tag{2.77}
$$

Proof: Consider functions T and R with

$$E \xrightarrow{T} F \xrightarrow{R} H$$

for some spaces E, F and H, where

$$\begin{aligned} T(v) &= T_1(v) + \frac{1}{2}T_2(v,v) + \frac{1}{6}T_3(v,v,v) + \ldots \\ R(w) &= R_1(w) + \frac{1}{2}R_2(w,w) + \frac{1}{6}R_3(w,w,w) + \ldots \end{aligned}$$

for $v \in E$ and $w \in F$. We can obtain a composition rule

$$\begin{aligned} R(T(v)) &= R_1(T_1(v) + \frac{1}{2}T_2(v,v) + \frac{1}{6}T_3(v,v,v) + \ldots) \\ &\quad + \frac{1}{2}R_2(T_1(v) + \frac{1}{2}T_2(v,v) + \ldots, T_1(v) + \frac{1}{2}T_2(v,v) + \ldots) \\ &\quad + \frac{1}{6}R_3(T_1(v) + \ldots, T_1(v) + \ldots, T_1(v) + \ldots) + O(v^4) \\ &= R_1(T_1(v)) + \frac{1}{2}\left[R_1(T_2(v,v)) + R_2(T_1(v), T_1(v))\right] \\ &\quad + \frac{1}{6}\left[R_1(T_3(v,v,v)) + 3R_2(T_1(v), T_2(v,v)) + R_3(T_1(v), T_1(v), T_1(v))\right] \\ &\quad + O(v^4). \end{aligned} \tag{2.78}$$

We may use the above in the case where $H = E$ to find the inverse of T, by putting

$$R(T(v)) = v,$$

and matching terms of each order in v. Thus, considering terms of each order in turn, we have

$O(v)$:

$$\begin{aligned} R_1(T_1(v)) &= v \\ \Rightarrow \quad R_1(w) &= T_1^{-1}(w), \end{aligned} \tag{2.79}$$

$O(v^2)$:

$$\begin{aligned} R_1(T_2(v,v)) + R_2(T_1(v), T_1(v)) &= 0 \\ \Rightarrow R_2(w,w) &= -T_1^{-1}(T_2(T_1^{-1}(w), T_1^{-1}(w))), \end{aligned} \tag{2.80}$$

$O(v^3)$:

$$\begin{aligned} &R_1(T_3(v,v,v)) + 3R_2(T_1(v), T_2(v,v)) + R_3(T_1(v), T_1(v), T_1(v)) = 0, \\ \Rightarrow \quad &R_3(w,w,w) = -T_1^{-1}(T_3(T_1^{-1}(w), T_1^{-1}(w), T_1^{-1}(w))) \\ &\qquad + 3T_1^{-1}(T_2(T_1^{-1}(w), T_1^{-1}(T_2(T_1^{-1}(w), T_1^{-1}(w)))))). \end{aligned} \tag{2.81}$$

To apply Lemma 2.1, we consider $\tilde{\delta}^{\xi}$ and Z_{ξ} as vectors. Using the notation $[T_1(\tilde{\delta}^{\xi})]_i$ for the ith component of the vector-valued function T_1 and so on, and comparing (2.72) with (2.73), we have

$$\begin{aligned}
[T_1(\tilde{\delta}^{\xi})]_{\xi_1} &= \left(\nu_{\xi_1\xi_2} + n^{-\frac{1}{2}} Z_{\xi_1\xi_2}\right)\tilde{\delta}^{\xi_2} && (2.82)\\
[T_2(\tilde{\delta}^{\xi}, \tilde{\delta}^{\xi})]_{\xi_1} &= \left(n^{-\frac{1}{2}}\nu_{\xi_1\xi_2\xi_3} + n^{-1} Z_{\xi_1\xi_2\xi_3}\right)\tilde{\delta}^{\xi_2}\tilde{\delta}^{\xi_3} && (2.83)\\
[T_3(\tilde{\delta}^{\xi}, \tilde{\delta}^{\xi}, \tilde{\delta}^{\xi})]_{\xi_1} &= n^{-1}\nu_{\xi_1\xi_2\xi_3\xi_4}\tilde{\delta}^{\xi_2}\tilde{\delta}^{\xi_3}\tilde{\delta}^{\xi_4}. && (2.84)
\end{aligned}$$

T_1^{-1} is the matrix inverse of $\nu_{\xi\xi} + n^{-\frac{1}{2}} Z_{\xi\xi}$. Thus, using (2.58) we have

$$\begin{aligned}
[R_1(Z_\xi)]^{\xi_1} &= -b^{\xi_1\xi_2} Z_{\xi_2}\\
&\quad + n^{-\frac{1}{2}} b^{\xi_1\xi_2} Z_{\xi_2\xi_3} b^{\xi_3\xi_4} Z_{\xi_4}\\
&\quad - n^{-1} b^{\xi_1\xi_2} Z_{\xi_2\xi_3} b^{\xi_3\xi_4} Z_{\xi_4\xi_5} b^{\xi_5\xi_6} Z_{\xi_6}\\
&\quad + O_p(n^{-\frac{3}{2}}) && (2.85)\\
[R_2(Z_\xi, Z_\xi)]^{\xi_1} &= -n^{-\frac{1}{2}} b^{\xi_1\xi_2}\nu_{\xi_2\xi_3\xi_4} b^{\xi_3\xi_5} Z_{\xi_5} b^{\xi_4\xi_6} Z_{\xi_6}\\
&\quad - n^{-1}\Big\{ b^{\xi_1\xi_2} Z_{\xi_2\xi_3\xi_4} b^{\xi_3\xi_5} Z_{\xi_5} b^{\xi_4\xi_6} Z_{\xi_6}\\
&\qquad - 2 b^{\xi_1\xi_2}\nu_{\xi_2\xi_3\xi_4} b^{\xi_3\xi_5} Z_{\xi_5} b^{\xi_4\xi_6} Z_{\xi_6\xi_7} b^{\xi_7\xi_8} Z_{\xi_8}\\
&\qquad - b^{\xi_1\xi_2} Z_{\xi_2\xi_3} b^{\xi_3\xi_4}\nu_{\xi_4\xi_5\xi_6} b^{\xi_5\xi_7} Z_{\xi_7} b^{\xi_6\xi_8} Z_{\xi_8}\Big\}\\
&\quad + O_p(n^{-\frac{3}{2}}) && (2.86)\\
[R_3(Z_\xi, Z_\xi, Z_\xi)]^{\xi_1} &= n^{-1}\Big\{ b^{\xi_1\xi_2}\nu_{\xi_2\xi_3\xi_4\xi_5} b^{\xi_3\xi_6} Z_{\xi_6} b^{\xi_4\xi_7} Z_{\xi_7} b^{\xi_5\xi_8} Z_{\xi_8}\\
&\qquad - 3 b^{\xi_1\xi_2}\nu_{\xi_2\xi_3\xi_4} b^{\xi_3\xi_5} Z_{\xi_5} b^{\xi_4\xi_6}\\
&\qquad\qquad \nu_{\xi_6\xi_7\xi_8} b^{\xi_7\xi_9} Z_{\xi_9} b^{\xi_8\xi_{10}} Z_{\xi_{10}}\Big\}\\
&\quad + O_p(n^{-\frac{3}{2}}), && (2.87)
\end{aligned}$$

where $b^{\xi_1\xi_2}$ is the (ξ_1, ξ_2) entry of the partitioned matrix

$$B = \begin{bmatrix} 0 & 0 \\ 0 & H_{\xi\xi}^{-1} \end{bmatrix}. \tag{2.88}$$

Hence we may write $\tilde{\delta}^{\xi_1}$ explicitly in terms of the Zs as

$$\begin{aligned}
\tilde{\delta}^{\xi_1} &= -b^{\xi_1\xi_2} Z_{\xi_2}\\
&\quad + n^{-\frac{1}{2}}\left\{ b^{\xi_1\xi_2} Z_{\xi_2\xi_3} b^{\xi_3\xi_4} Z_{\xi_4} - \frac{1}{2} b^{\xi_1\xi_2}\nu_{\xi_2\xi_3\xi_4} b^{\xi_3\xi_5} Z_{\xi_5} b^{\xi_4\xi_6} Z_{\xi_6}\right\}\\
&\quad + n^{-1}\Big\{ - b^{\xi_1\xi_2} Z_{\xi_2\xi_3} b^{\xi_3\xi_4} Z_{\xi_4\xi_5} b^{\xi_5\xi_6} Z_{\xi_6}
\end{aligned}$$

$$
\begin{aligned}
&\quad -\frac{1}{2}b^{\xi_1\xi_2}Z_{\xi_2\xi_3\xi_4}b^{\xi_3\xi_5}Z_{\xi_5}b^{\xi_4\xi_6}Z_{\xi_6}\\
&\quad +b^{\xi_1\xi_2}\nu_{\xi_2\xi_3\xi_4}b^{\xi_3\xi_5}Z_{\xi_5}b^{\xi_4\xi_6}Z_{\xi_6\xi_7}b^{\xi_7\xi_8}Z_{\xi_8}\\
&\quad +\frac{1}{2}b^{\xi_1\xi_2}Z_{\xi_2\xi_3}b^{\xi_3\xi_4}\nu_{\xi_4\xi_5\xi_6}b^{\xi_5\xi_7}Z_{\xi_7}b^{\xi_6\xi_8}Z_{\xi_8}\\
&\quad +\frac{1}{6}b^{\xi_1\xi_2}\nu_{\xi_2\xi_3\xi_4\xi_5}b^{\xi_3\xi_6}Z_{\xi_6}b^{\xi_4\xi_7}Z_{\xi_7}b^{\xi_6\xi_8}Z_{\xi_8}\\
&\quad -\frac{1}{2}b^{\xi_1\xi_2}\nu_{\xi_2\xi_3\xi_4}b^{\xi_3\xi_5}Z_{\xi_5}b^{\xi_4\xi_8}\nu_{\xi_8\xi_9\xi_{10}}b^{\xi_9\xi_{11}}Z_{\xi_{11}}b^{\xi_{10}\xi_{12}}Z_{\xi_{12}}\Big\}\\
&+O(n^{-\frac{3}{2}}).
\end{aligned}
\tag{2.89}
$$

Now using (2.71) in (2.36) we have

$$
S_R=\tilde{Z}_{\tau_1}\tilde{\nu}^{\tau_1,\tau_2}\tilde{Z}_{\tau_2},
\tag{2.90}
$$

and substituting (2.89) into (2.72) and (2.70) we have

$$
\begin{aligned}
\tilde{Z}_{\tau_1} \;=\;& Z_{\tau_1}-\nu_{\tau_1\xi_1}b^{\xi_1\xi_2}Z_{\xi_2}\\
&+n^{-\frac{1}{2}}\Big\{\left(\nu_{\tau_1\xi_1}b^{\xi_1\xi_2}Z_{\xi_2\xi_3}-Z_{\tau_1\xi_3}\right)b^{\xi_3\xi_4}Z_{\xi_4}\\
&\qquad -\frac{1}{2}\left(\nu_{\tau_1\xi_1}b^{\xi_1\xi_2}\nu_{\xi_2\xi_3\xi_4}-\nu_{\tau_1\xi_3\xi_4}\right)b^{\xi_3\xi_5}Z_{\xi_5}b^{\xi_4\xi_6}Z_{\xi_6}\Big\}\\
&+n^{-1}\Big\{\left(Z_{\tau_1\xi_1}-\nu_{\tau_1\xi_2}b^{\xi_2\xi_3}Z_{\xi_3\xi_1}\right)b^{\xi_1\xi_4}Z_{\xi_4\xi_5}b^{\xi_5\xi_6}Z_{\xi_6}\\
&\qquad +\left(\nu_{\tau_1\xi_1}b^{\xi_1\xi_2}\nu_{\xi_2\xi_3\xi_4}-\nu_{\tau_1\xi_3\xi_4}\right)b^{\xi_3\xi_5}Z_{\xi_5}b^{\xi_4\xi_6}Z_{\xi_6\xi_7}b^{\xi_7\xi_8}Z_{\xi_8}\\
&\qquad -\frac{1}{2}\left(\nu_{\tau_1\xi_1}b^{\xi_1\xi_2}Z_{\xi_2\xi_3\xi_4}-Z_{\tau_1\xi_3\xi_4}\right)b^{\xi_3\xi_5}Z_{\xi_5}b^{\xi_4\xi_6}Z_{\xi_6}\\
&\qquad +\frac{1}{6}\left(\nu_{\tau_1\xi_1}b^{\xi_1\xi_2}\nu_{\xi_2\xi_3\xi_4\xi_5}-\nu_{\tau_1\xi_3\xi_4\xi_5}\right)b^{\xi_3\xi_6}Z_{\xi_6}b^{\xi_4\xi_7}Z_{\xi_7}b^{\xi_5\xi_8}Z_{\xi_8}\\
&\qquad -\frac{1}{2}\left(\nu_{\tau_1\xi_1}b^{\xi_1\xi_2}\nu_{\xi_2\xi_3\xi_4}-\nu_{\tau_1\xi_3\xi_4}\right)b^{\xi_3\xi_5}Z_{\xi_5}b^{\xi_4\xi_6}\nu_{\xi_6\xi_7\xi_8}b^{\xi_7\xi_9}Z_{\xi_9}b^{\xi_8\xi_{10}}Z_{\xi_{10}}\\
&\qquad +\frac{1}{2}\left(\nu_{\tau_1\xi_1}b^{\xi_1\xi_2}Z_{\xi_2\xi_3}-Z_{\tau_1\xi_3}\right)b^{\xi_3\xi_4}\nu_{\xi_4\xi_5\xi_6}b^{\xi_5\xi_7}Z_{\xi_7}b^{\xi_6\xi_8}\Big\}\\
&+O(n^{-\frac{3}{2}})
\end{aligned}
\tag{2.91}
$$

and

$$
\begin{aligned}
\tilde{\nu}^{\tau_1,\tau_2} \;=\;& \nu^{\tau_1,\tau_2}\\
&+n^{-\frac{1}{2}}\nu^{\tau_1,\phi_1}\nu_{\phi_1,\phi_2,\xi_1}b^{\xi_1\xi_2}Z_{\xi_2}\nu^{\phi_2,\tau_2}\\
&+n^{-1}\Big\{-\nu^{\tau_1,\phi_1}\nu_{\phi_1,\phi_2,\xi_1}\\
&\qquad\qquad (b^{\xi_1\xi_2}Z_{\xi_2\xi_3}b^{\xi_3\xi_4}Z_{\xi_4}-\frac{1}{2}b^{\xi_1\xi_2}\nu_{\xi_2\xi_3\xi_4}b^{\xi_3\xi_5}Z_{\xi_5}b^{\xi_4\xi_6}Z_{\xi_6})\\
&\qquad +\frac{1}{2}\nu^{\tau_1,\phi_1}(\nu_{\phi_1,\phi_2,\xi_1\xi_2}+\nu_{\phi_1,\xi_1,\phi_2\xi_2}+\nu_{\phi_2,\xi_1,\phi_1\xi_2}
\end{aligned}
$$

$$
\begin{aligned}
& \qquad + \nu_{\phi_1,\phi_2,\xi_1,\xi_2}) b^{\xi_1\xi_3} Z_{\xi_3} b^{\xi_2\xi_4} Z_{\xi_4} \nu^{\phi_2,\tau_2} \\
& \qquad + \nu^{\tau_1,\phi_1} \nu_{\phi_1,\phi_2,\xi_1} b^{\xi_1\xi_2} Z_{\xi_2} \nu^{\phi_2,\phi_3} \nu_{\phi_3,\phi_4,\xi_3} b^{\xi_3\xi_4} Z_{\xi_4} \nu^{\phi_4,\tau_2} \Big\} \\
& + O(n^{-\frac{3}{2}}).
\end{aligned} \tag{2.92}
$$

Note that the first order term in the expansion for $\tilde{Z}_\tau$ is of the form

$$
Z_\tau - K_{\tau\xi} A^{\xi\xi} Z_\xi, \tag{2.93}
$$

which we may regard as a "horizontal" component of Z_ϕ. To see the reason for this interpretation consider

$$
\begin{aligned}
\mathrm{E}\left[\left(Z_{\tau_1} - \nu_{\tau_1\xi_1} b^{\xi_1\xi_2} Z_{\xi_2}\right) Z_{\xi_3}\right] &= \nu_{\tau_1,\xi_3} - \nu_{\tau_1\xi_1} b^{\xi_1\xi_2} \nu_{\xi_2,\xi_3} \\
&= \nu_{\tau_1,\xi_3} - \nu_{\tau_1,\xi_1} a^{\xi_1\xi_2} \nu_{\xi_2,\xi_3},
\end{aligned}
$$

where $a^{\xi_1\xi_2}$ is the (ξ_1, ξ_2) entry of the partitioned matrix

$$
A = \begin{bmatrix} 0 & 0 \\ 0 & K_{\xi,\xi}^{-1} \end{bmatrix}
$$

as before (see (2.50)). Hence

$$
\begin{aligned}
\mathrm{E}\left[\left(Z_{\tau_1} - \nu_{\tau_1\xi_1} b^{\xi_1\xi_2} Z_{\xi_2}\right) Z_{\xi_3}\right] &= \nu_{\tau_1,\xi_3} - \nu_{\tau_1,\xi_1} \delta_{\xi_3}^{\xi_1} \\
&= 0,
\end{aligned} \tag{2.94}
$$

i.e. the first order term of $\tilde{Z}_\tau$ is orthogonal to Z_ξ, the "vertical" component of the score Z_ϕ. The higher order terms in the expansion for $\tilde{Z}_\tau$ contain analogous expressions, so we introduce a more concise notation by writing

$$
\begin{aligned}
Z_{h_1} &= Z_{\tau_1} - \nu_{\tau_1\xi_1} b^{\xi_1\xi_2} Z_{\xi_2} && (2.95) \\
Z_{h_1\xi_1} &= Z_{\tau_1\xi_1} - \nu_{\tau_1\xi_2} b^{\xi_2\xi_3} Z_{\xi_3\xi_1} && (2.96) \\
\nu_{h_1\xi_1\xi_2} &= \nu_{\tau_1\xi_1\xi_2} - \nu_{\tau_1\xi_3} b^{\xi_3\xi_4} \nu_{\xi_4\xi_1\xi_2}, && (2.97)
\end{aligned}
$$

and then, in order to maintain the pairing of indices in the index notation, we introduce the convention that h as a superscript denotes summation over the range $1, \ldots, p$, i.e. over the range of indices corresponding to the interest parameter. We may therefore write, for example, the first order term of S_R as

$$
Z_{h_1} \nu^{h_1h_2} Z_{h_2} = \left(Z_{\tau_1} - \nu_{\tau_1\xi_1} b^{\xi_1\xi_2} Z_{\xi_2}\right) \nu^{\tau_1\tau_2} \left(Z_{\tau_2} - \nu_{\tau_2\xi_3} b^{\xi_3\xi_4} Z_{\xi_4}\right). \tag{2.98}
$$

Combining (2.91) and (2.92), the full expansion for S_R may be written as

$$S_R = S_0 + n^{-\frac{1}{2}} S_1 + n^{-1} S_2 + O(n^{-\frac{3}{2}}), \tag{2.99}$$

where

$$\begin{aligned}
S_0 &= \nu^{h_1,h_2} Z_{h_1} Z_{h_2} &(2.100)\\
S_1 &= +\nu^{h_1,\phi_1}\nu_{\xi_1,\phi_1,\phi_2} b^{\xi_1\xi_2}\nu^{\phi_2,h_2} Z_{\xi_2} Z_{h_1} Z_{h_2}\\
&\quad -2\nu^{h_1,h_2} b^{\xi_1\xi_2} Z_{h_1\xi_1} Z_{\xi_2} Z_{h_2}\\
&\quad +\nu^{h_1,h_2} b^{\xi_1\xi_3} b^{\xi_2\xi_4}\nu_{h_1\xi_1\xi_2} Z_{\xi_3} Z_{\xi_4} Z_{h_2} &(2.101)\\
S_2 &= -\nu^{h_1,\phi_1}\nu_{\xi_1,\phi_1,\phi_2} b^{\xi_1\xi_2} b^{\xi_3\xi_4}\nu^{\phi_2,h_2} Z_{h_1} Z_{h_2} Z_{\xi_2\xi_3} Z_{\xi_4}\\
&\quad +\frac{1}{2}\nu^{h_1,\phi_1}\nu_{\xi_1,\phi_1,\phi_2} b^{\xi_1\xi_2}\nu_{\xi_2\xi_3\xi_4} b^{\xi_3\xi_5} b^{\xi_4\xi_6}\nu^{\phi_2,h_2} Z_{\xi_6} Z_{\xi_5} Z_{h_1} Z_{h_2}\\
&\quad +\frac{1}{2}\nu^{h_1,\phi_1}\left(\nu_{\phi_1,\phi_2,\xi_1\xi_2} + 2\nu_{\phi_1,\xi_1,\phi_2\xi_2} + \nu_{\phi_1,\phi_2,\xi_1,\xi_2}\right) b^{\xi_1\xi_3} b^{\xi_2\xi_4}\nu^{\phi_2,h_2}\\
&\qquad\qquad Z_{\xi_3} Z_{\xi_4} Z_{h_1} Z_{h_2}\\
&\quad +\nu^{h_1,\phi_1}\nu_{\xi_1,\phi_1,\phi_2} b^{\xi_1\xi_2}\nu^{\phi_2\phi_3}\nu_{\xi_3,\phi_3,\phi_4} b^{\xi_3\xi_4}\nu^{\phi_4,h_2} Z_{\xi_2} Z_{\xi_4} Z_{h_1} Z_{h_2}\\
&\quad +2\nu^{h_1,h_2} b^{\xi_1\xi_2} b^{\xi_3\xi_4} Z_{h_1} Z_{h_2\xi_1} Z_{\xi_2\xi_3} Z_{\xi_4}\\
&\quad -2\nu^{h_1,h_2}\nu_{h_2\xi_1\xi_2} b^{\xi_1\xi_3} b^{\xi_2\xi_4} b^{\xi_5\xi_6} Z_{h_1} Z_{\xi_3} Z_{\xi_4\xi_5} Z_{\xi_6}\\
&\quad +\nu^{h_1,h_2} b^{\xi_1\xi_2} b^{\xi_3\xi_4} Z_{h_1} Z_{h_2\xi_1\xi_3} Z_{\xi_2} Z_{\xi_4}\\
&\quad -\frac{1}{3}\nu^{h_1,h_2}\nu_{h_2\xi_1\xi_2\xi_3} b^{\xi_1\xi_4} b^{\xi_2\xi_5} b^{\xi_3\xi_6} Z_{h_1} Z_{\xi_4} Z_{\xi_5} Z_{\xi_6}\\
&\quad +\nu^{h_1,h_2}\nu_{h_2\xi_1\xi_2} b^{\xi_1\xi_3} b^{\xi_2\xi_4}\nu_{\xi_4\xi_5\xi_6} b^{\xi_5\xi_7} b^{\xi_6\xi_8} Z_{h_1} Z_{\xi_3} Z_{\xi_7} Z_{\xi_8}\\
&\quad -\nu^{h_1,h_2} b^{\xi_1\xi_2}\nu_{\xi_2\xi_3\xi_4} b^{\xi_3\xi_5} b^{\xi_4\xi_6} Z_{h_1} Z_{h_2\xi_1} Z_{\xi_5} Z_{\xi_6}\\
&\quad +\nu^{h_1,h_2} b^{\xi_1\xi_2} b^{\xi_3\xi_4} Z_{h_1\xi_1} Z_{\xi_2} Z_{h_2\xi_3} Z_{\xi_4}\\
&\quad -\nu^{h_1,h_2} b^{\xi_1\xi_2}\nu_{h_2\xi_3\xi_4} b^{\xi_3\xi_5} b^{\xi_4\xi_6} Z_{h_1\xi_1} Z_{\xi_2} Z_{\xi_5} Z_{\xi_6}\\
&\quad +\frac{1}{4}\nu^{h_1,h_2}\nu_{h_1\xi_1\xi_2} b^{\xi_1\xi_3} b^{\xi_2\xi_4}\nu_{h_2\xi_5\xi_6} b^{\xi_5\xi_7} b^{\xi_6\xi_8} Z_{\xi_3} Z_{\xi_4} Z_{\xi_7} Z_{\xi_8}\\
&\quad -2\nu^{h_1,\phi_1}\nu_{\xi_1,\phi_1,\phi_2} b^{\xi_1\xi_2}\nu^{\phi_2 h_2} b^{\xi_3\xi_4} Z_{\xi_2} Z_{h_1} Z_{h_2\xi_3} Z_{\xi_4}\\
&\quad +\nu^{h_1,\phi_1}\nu_{\xi_1,\phi_1,\phi_2} b^{\xi_1\xi_2}\nu^{\phi_2,h_2}\nu_{h_2\xi_3\xi_4} b^{\xi_3\xi_5} b^{\xi_4\xi_6} Z_{\xi_2} Z_{h_1} Z_{\xi_5} Z_{\xi_6}. &(2.102)
\end{aligned}$$

2.2.7 The Moment Generating Function of S_R

We now wish to obtain from this expansion of S_R an asymptotic expansion for the moment generating function, $M_{S_R}(t)$, of S_R, in order to calculate the the correct expressions for A_1, A_2 and A_3. The only random variables in the expressions (2.100)

- (2.102) for the components of S_R are the Zs. We may define a vector $\mathbf{Z}$ as

$$\begin{aligned}
\mathbf{Z} &= (Z_\phi, Z_{\phi,\phi}, Z_{\phi,\phi,\phi}) \\
&= (Z_1, \ldots, Z_m, \\
&\quad Z_{11}, \ldots, Z_{1m}, \ldots, Z_{m1}, \ldots, Z_{mm}, \\
&\quad Z_{111}, \ldots, Z_{11m}, \ldots, Z_{mm1}, \ldots, Z_{mmm}). \qquad (2.103)
\end{aligned}$$

In order to find

$$M_{S_R}(t) = \mathrm{E}\left[e^{S_R t}\right],$$

we find an approximation to $f_{\mathbf{Z}}(x)$, the density function of $\mathbf{Z}$ up to order $O(n^{-1})$.

The Edgeworth expansion for the density function of a variable X, having its rth cumulant of order $O(n^{1-\frac{r}{2}})$, is

$$f_X(x) = \phi(x;\eta)\left\{1 + Q_3\left(x;\eta^{(3)}\right) + Q_4\left(x;\eta^{(4)}\right)\right\} + O\left(n^{-\frac{3}{2}}\right) \qquad (2.104)$$

(see, for example, Barndorff-Nielsen, Blæsild, Pace & Salvan (1991, section 12.1)). Here η is the covariance matrix of X, $\eta^{(3)}$ and $\eta^{(4)}$ are the third and fourth order cumulants of X, having elements $\eta^{i,j,k}$ and $\eta^{i,j,k,l}$ respectively, and Q_3 and Q_4 are given by

$$\begin{aligned}
Q_3\left(x;\eta^{(3)}\right) &= \frac{1}{6}\eta^{i,j,k}h_{ijk}(x;\eta) \qquad (2.105)\\
Q_4\left(x;\eta^{(4)}\right) &= \frac{1}{24}\eta^{i,j,k,l}h_{ijkl}(x;\eta) + \frac{1}{72}\eta^{i,j,k}\eta^{l,m,n}h_{ijklmn}(x;\eta). \qquad (2.106)
\end{aligned}$$

Summation is, of course, over all of the terms in X, i.e. from 1 to d if X has dimension d. The terms h_{ijk}, h_{ijkl} and h_{ijklmn} are covariant Hermite polynomials given by

$$\begin{aligned}
h_{ijk}(x;\eta) &= x_i x_j x_k - x_i \eta_{j,k}[3], \qquad (2.107)\\
h_{ijkl}(x;\eta) &= x_i x_j x_k x_l - x_i x_j \eta_{k,l}[6] + \eta_{i,j}\eta_{k,l}[3], \qquad (2.108)\\
h_{ijklmn}(x;\eta) &= x_i x_j x_k x_l x_m x_n - x_i x_j x_k x_l \eta_{m,n}[15] \\
&\quad + x_i x_j \eta_{k,l}\eta_{m,n}[45] - \eta_{i,j}\eta_{k,l}\eta_{m,n}[15], \qquad (2.109)
\end{aligned}$$

where $\eta_{i,j}$ is the (i,j)-element of the inverse of the covariance matrix η and $x_i = \eta_{i,j}x^j$. Clearly $Q_3(x;\eta^{(3)})$ is of order $O(n^{-\frac{1}{2}})$, and $Q_4(x;\eta^{(4)})$ is of order $O(n^{-1})$. Note that here we replace the x^i by Z_i, Z_{ij}, etc., so we use the version where Q_3

and Q_4 are defined in terms of the contravariant Hermite polynomials, i.e.

$$\begin{aligned} Q_3\left(x;\eta_{(3)}\right) &= \frac{1}{6}\eta_{i,j,k}h^{ijk}(x;\eta), & (2.110)\\ h^{ijk}(x;\eta) &= x^i x^j x^k - x^i\eta^{j,k}[3], & (2.111)\end{aligned}$$

and so on, with $x^i = \eta^{i,j}x_j$.

Now we write the covariance matrix of $\mathbf{Z}$ as

$$V = \begin{bmatrix} V_{11} & V_{12} & V_{13} \\ V_{21} & V_{22} & V_{23} \\ V_{31} & V_{32} & V_{33} \end{bmatrix}, \tag{2.112}$$

where V_{11} is an $m \times m$ submatrix, V_{12} is an $m \times m^2$ submatrix, and so on. In the usual way, we write the inverse as

$$V^{-1} = \begin{bmatrix} V^{11} & V^{12} & V^{13} \\ V^{21} & V^{22} & V^{23} \\ V^{31} & V^{32} & V^{33} \end{bmatrix}. \tag{2.113}$$

We will also write $\mathbf{z}$ for the observed value of $\mathbf{Z}$ and $\nu_{(3)}$ and $\nu_{(4)}$ for the third and fourth order cumulants of $\mathbf{Z}$ respectively. Note that the matrix V considered in more detail is

$$\begin{aligned} V &= \left[\begin{array}{ccc|ccc|ccc} \nu_{1,1} & \cdots & \nu_{1,m} & \nu_{1,11} & \cdots & \nu_{1,mm} & \nu_{1,111} & \cdots & \nu_{1,mmm} \\ \vdots & & \vdots & \vdots & & \vdots & \vdots & & \vdots \\ \nu_{m,1} & \cdots & \nu_{m,m} & \nu_{m,11} & \cdots & \nu_{m,mm} & \nu_{m,111} & \cdots & \nu_{m,mmm} \\ \hline \nu_{11,1} & \cdots & \nu_{11,m} & \nu_{11,11} & \cdots & \nu_{11,mm} & \nu_{11,111} & \cdots & \nu_{11,mmm} \\ \vdots & & \vdots & \vdots & & \vdots & \vdots & & \vdots \\ \nu_{mm,1} & \cdots & \nu_{mm,m} & \nu_{mm,11} & \cdots & \nu_{mm,mm} & \nu_{mm,111} & \cdots & \nu_{mm,mmm} \\ \hline \nu_{111,1} & \cdots & \nu_{111,m} & \nu_{111,11} & \cdots & \nu_{111,mm} & \nu_{111,111} & \cdots & \nu_{111,mmm} \\ \vdots & & \vdots & \vdots & & \vdots & \vdots & & \vdots \\ \nu_{mmm,1} & \cdots & \nu_{mmm,m} & \nu_{mmm,11} & \cdots & \nu_{mmm,mm} & \nu_{mmm,111} & \cdots & \nu_{mmm,mmm} \end{array}\right] \\ &= \begin{bmatrix} V_{11} & 0 & 0 \\ 0 & V_{22} & V_{23} \\ 0 & V_{32} & V_{33} \end{bmatrix}, \end{aligned} \tag{2.114}$$

by equations (2.25) and (2.26).

Using (2.104) we have, integrating over all possible values of $\mathbf{Z}$,

$$\begin{aligned} M_{S_R}(t) &= \int e^{S_R t} \mid 2\pi V \mid^{-\frac{1}{2}} e^{-\frac{1}{2}\mathbf{z}^T V^{-1}\mathbf{z}} \left\{1 + Q_3\left(\mathbf{z};\nu_{(3)}\right) + Q_4\left(\mathbf{z};\nu_{(4)}\right)\right\} d\mathbf{z} \\ &\quad + O\left(n^{-\frac{3}{2}}\right) \\ &= \int \mid 2\pi V \mid^{-\frac{1}{2}} e^{S_0(\mathbf{z})t - \frac{1}{2}\mathbf{z}^T V^{-1}\mathbf{z}} e^{\left(n^{-\frac{1}{2}}S_1(\mathbf{z}) + n^{-1}S_2(\mathbf{z})\right)t} \end{aligned}$$

$$
\begin{aligned}
& \left\{1+Q_3\left(\mathbf{z};\nu_{(3)}\right)+Q_4\left(\mathbf{z};\nu_{(4)}\right)\right\}\,d\mathbf{z}+O\left(n^{-\frac{3}{2}}\right) \\
= & \int \mid 2\pi V \mid^{-\frac{1}{2}} e^{S_0(\mathbf{z})t-\frac{1}{2}\mathbf{z}^T V^{-1}\mathbf{z}}\left\{1+n^{-\frac{1}{2}}S_1(\mathbf{z})t\right. \\
& \left. +n^{-1}\left[\frac{1}{2}(S_1(\mathbf{z}))^2t^2+S_2(\mathbf{z})t\right]\right\}\left\{1+Q_3\left(\mathbf{z};\nu_{(3)}\right)+Q_4\left(\mathbf{z};\nu_{(4)}\right)\right\}\,d\mathbf{z} \\
& \qquad +O\left(n^{-\frac{3}{2}}\right) \\
= & \int \mid 2\pi V \mid^{-\frac{1}{2}} e^{S_0(\mathbf{z})t-\frac{1}{2}\mathbf{z}^T V^{-1}\mathbf{z}} \\
& \qquad\qquad \left\{1+\left[n^{-\frac{1}{2}}S_1(\mathbf{z})t+Q_3\left(\mathbf{z};\nu_{(3)}\right)\right]\right. \\
& \qquad\qquad\quad +\left[n^{-1}\frac{1}{2}(S_1(\mathbf{z}))^2t^2+n^{-1}S_2(\mathbf{z})t\right. \\
& \qquad\qquad\qquad \left.\left. +n^{-\frac{1}{2}}S_1(\mathbf{z})Q_3\left(\mathbf{z};\nu_{(3)}\right)t+Q_4\left(\mathbf{z};\nu_{(4)}\right)\right]\right\}\,d\mathbf{z} \\
& +O\left(n^{-\frac{3}{2}}\right),
\end{aligned} \tag{2.115}
$$

where S_0, S_1 and S_2 have been written explicitly as functions of $\mathbf{z}$ for clarity. Now

$$
S_0(\mathbf{z})=\left(Z_{\tau_1}-\nu_{\tau_1\xi_1}b^{\xi_1\xi_2}Z_{\xi_2}\right)\nu^{\tau_1,\tau_2}\left(Z_{\tau_2}-\nu_{\tau_2\xi_3}b^{\xi_3\xi_4}Z_{\xi_4}\right),
$$

in index notation, which we may write in matrix notation as

$$
\begin{aligned}
S_0(\mathbf{z}) &= \left[\begin{array}{cc} Z_\tau^T & Z_\xi^T \end{array}\right]\left[\begin{array}{cc} K^{\tau,\tau} & -K^{\tau,\tau}H_{\tau\xi}H_{\xi\xi}^{-1} \\ -H_{\xi\xi}^{-1}H_{\xi\tau}K^{\tau,\tau} & H_{\xi\xi}^{-1}H_{\xi\tau}K^{\tau,\tau}H_{\tau\xi}H_{\xi\xi}^{-1} \end{array}\right]\left[\begin{array}{c} Z_\tau \\ Z_\xi \end{array}\right] \\
&= \left[\begin{array}{cc} Z_\tau^T & Z_\xi^T \end{array}\right] M \left[\begin{array}{c} Z_\tau \\ Z_\xi \end{array}\right],
\end{aligned} \tag{2.116}
$$

where M is defined in (2.51). Now if we define a partitioned matrix U by

$$
U=\left[\begin{array}{ccc} M & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{array}\right] \tag{2.117}
$$

then we may write (2.115) as

$$
\begin{aligned}
M_{S_R}(t) = & \frac{\mid 2\pi V \mid^{-\frac{1}{2}}}{\mid 2\pi W \mid^{-\frac{1}{2}}}\int \mid 2\pi W \mid^{-\frac{1}{2}} e^{-\frac{1}{2}\mathbf{z}^T W^{-1}\mathbf{z}} \\
& \qquad \left\{1+\left[n^{-\frac{1}{2}}S_1(\mathbf{z})t+Q_3\left(\mathbf{z};\nu_{(3)}\right)\right]\right. \\
& \qquad\quad +\left[n^{-1}\frac{1}{2}(S_1(\mathbf{z}))^2t^2+n^{-1}S_2(\mathbf{z})t\right. \\
& \qquad\quad \left.\left. +n^{-\frac{1}{2}}S_1(\mathbf{z})Q_3\left(\mathbf{z};\nu_{(3)}\right)t+Q_4\left(\mathbf{z};\nu_{(4)}\right)\right]\right\}\,d\mathbf{z} \\
& +O\left(n^{-\frac{3}{2}}\right),
\end{aligned} \tag{2.118}
$$

where the matrix W^{-1} is given by

$$
\begin{aligned}
W^{-1} &= V^{-1}-2tU \\
&= V^{-1}(I-2tVU).
\end{aligned} \tag{2.119}
$$

Thus to find $M_{S_n}(t)$ we must find the expected value of

$$\begin{aligned}P(\mathbf{Z}) &= 1+\left[n^{-\frac{1}{2}}S_1(\mathbf{Z})t+Q_3\left(\mathbf{Z};\nu_{(3)}\right)\right] \\ &\quad +\left[n^{-1}\frac{1}{2}(S_1(\mathbf{Z}))^2t^2+n^{-1}S_2(\mathbf{Z})t\right. \\ &\quad \left.+n^{-\frac{1}{2}}S_1(\mathbf{Z})Q_3\left(\mathbf{Z};\nu_{(3)}\right)t+Q_4\left(\mathbf{Z};\nu_{(4)}\right)\right] \qquad (2.120)\end{aligned}$$

where $\mathbf{Z}$ has a multivariate normal distribution with covariance matrix W.

Considering W now in more detail, we see that

$$\begin{aligned}W &= (I-2tVU)^{-1}V \\ &= \begin{bmatrix} I-2tV_{11}M & 0 & 0 \\ 0 & I & 0 \\ 0 & 0 & I\end{bmatrix}^{-1} V \\ &= \begin{bmatrix} (I-2tV_{11}M)^{-1} & 0 & 0 \\ 0 & I & 0 \\ 0 & 0 & I\end{bmatrix}\begin{bmatrix} V_{11} & 0 & 0 \\ 0 & V_{22} & V_{23} \\ 0 & V_{32} & V_{33}\end{bmatrix}. \qquad (2.121)\end{aligned}$$

Also

$$V_{11}M=\begin{bmatrix} K_{\tau,\tau} & K_{\tau,\xi} \\ K_{\xi,\tau} & K_{\xi,\xi}\end{bmatrix}\begin{bmatrix} K^{\tau,\tau} & -K^{\tau,\tau}H_{\tau\xi}H_{\xi\xi}^{-1} \\ -H_{\xi\xi}^{-1}H_{\xi\tau}K^{\tau,\tau} & H_{\xi\xi}^{-1}H_{\xi\tau}K^{\tau,\tau}H_{\tau\xi}H_{\xi\xi}^{-1}\end{bmatrix}, \qquad (2.122)$$

and if we now use the identities

$$\begin{aligned}H_{\tau\xi} &= -K_{\tau,\xi}, \\ H_{\xi\xi} &= -K_{\xi,\xi},\end{aligned}$$

and so on, we have

$$V_{11}M=\begin{bmatrix} I & -K_{\tau,\xi}K_{\xi,\xi}^{-1} \\ 0 & 0\end{bmatrix}. \qquad (2.123)$$

Clearly

$$(V_{11}M)^2=V_{11}M,$$

so, for small values of $|\ t\ |$, we have

$$\begin{aligned}(I-2tV_{11}M)^{-1} &= I+\sum_{r=1}^{\infty}(2t)^r(V_{11}M)^r \\ &= I+V_{11}M\sum_{r=1}^{\infty}(2t)^r \\ &= I+V_{11}M\frac{2t}{1-2t}, \qquad (2.124)\end{aligned}$$

thus we may write W as

$$W = V + \frac{2t}{1-2t} \begin{bmatrix} V_{11} M V_{11} & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}. \tag{2.125}$$

Substituting

$$\begin{aligned} \frac{\mid 2\pi V \mid^{-\frac{1}{2}}}{\mid 2\pi W \mid^{-\frac{1}{2}}} &= \frac{\mid 2\pi V \mid^{-\frac{1}{2}}}{\mid 2\pi V \mid^{-\frac{1}{2}} \mid (I - 2tVU)^{-1} \mid^{-\frac{1}{2}}} \\ &= \mid I - 2tVU \mid^{-\frac{1}{2}} \\ &= \mid I - 2tV_{11}M \mid^{-\frac{1}{2}} \\ &= \begin{vmatrix} (1-2t)I & 2tK_{\tau,\xi}K_{\xi,\xi}^{-1} \\ 0 & I \end{vmatrix}^{-\frac{1}{2}} \\ &= (1-2t)^{-\frac{p}{2}} \end{aligned} \tag{2.126}$$

into (2.118) we obtain

$$M_{S_R}(t) = \mathrm{E}\left[(1-2t)^{-\frac{p}{2}} P(\mathbf{Z})\right], \tag{2.127}$$

where $\mathbf{Z} \sim MVN(\mathbf{0}, W)$. Calculating $M_{S_R}(t)$ directly in this way is complicated - it is easier to proceed as follows.

Taking

$$d = \frac{2t}{1-2t},$$

we have, for $\mid t \mid < \frac{1}{2}$,

$$\begin{aligned} t &= \frac{d}{2(1+d)} \\ &= \frac{1}{2} \sum_{r=1}^{\infty} (-1)^{r+1} d^r \end{aligned} \tag{2.128}$$

and

$$t^2 = \frac{1}{4} \sum_{r=2}^{\infty} (-1)^r (r-1) d^r. \tag{2.129}$$

Substituting (2.128) and (2.129) into (2.120), we have

$$\begin{aligned} M_{S_R}(t) &= (1-2t)^{-\frac{p}{2}} \left\{ 1 + \mathrm{E}\left[n^{-\frac{1}{2}} \frac{1}{2} S_1(\mathbf{Z}) \sum_{r=1}^{\infty} (-1)^{r+1} d^r + Q_3\left(\mathbf{Z}; \nu_{(3)}\right) \right] \right. \\ &\quad + \mathrm{E}\left[n^{-1} \frac{1}{8} (S_1(\mathbf{Z}))^2 \sum_{r=2}^{\infty} (-1)^r (r-1) d^r \right. \\ &\quad + n^{-1} \frac{1}{2} S_2(\mathbf{Z}) \sum_{r=1}^{\infty} (-1)^{r+1} d^r \end{aligned}$$

$$+ n^{-\frac{1}{2}}\frac{1}{2}S_1(\mathbf{Z})Q_3\left(\mathbf{z};\nu_{(3)}\right)\sum_{r=1}^{\infty}(-1)^{r+1}d^r$$
$$+ Q_4\left(\mathbf{Z};\nu_{(4)}\right)\Big]\Big\}, \tag{2.130}$$

where, again, $\mathbf{Z} \sim MVN(\mathbf{0}, W)$. In general, for a variable $\mathbf{W} \sim MVN(\mathbf{0}, V)$, the even moments are

$$\begin{aligned}
\mathrm{E}(X_iX_j) &= V_{ij}, \\
\mathrm{E}(X_iX_jX_kX_l) &= V_{ij}V_{kl}[3], \\
\mathrm{E}(X_iX_jX_kX_lX_mX_n) &= V_{ij}V_{kl}V_{mn}[15],
\end{aligned}$$

and so on, and the odd moments are all zero. Thus the $O\left(n^{-\frac{1}{2}}\right)$ term of $M_{S_R}(t)$ is zero, and from (2.125) and (2.127) above we can see that the $O\left(n^{-1}\right)$ term of $M_{S_R}(t)$ is a polynomial in d. Using Harris's notation, and noting that $M_{S_R}(0) = 1$, we may write

$$M_{S_R}(t) = (1-2t)^{-\frac{p}{2}}\left\{1 + \frac{1}{24n}\left(A_1d + A_2d^2 + A_3d^3 + A_4d^4 + \ldots\right)\right\} + O\left(n^{-\frac{3}{2}}\right). \tag{2.131}$$

From this we find

$$\begin{aligned}
\mathrm{E}(S_R) &= p + \frac{1}{12n}A_1 && (2.132)\\
\mathrm{E}(S_R^2) &= p^2 + 2p + \frac{pA_1}{6n} + \frac{1}{3n}(A_1 + A_2) && (2.133)\\
\mathrm{E}(S_R^3) &= p^3 + 6p^2 + 8p + \frac{p^2A_1}{4n} + \frac{3pA_1}{2n} + \frac{pA_2}{n} \\
&\quad + \frac{1}{n}(A_1 + 4A_2 + 2A_3) && (2.134)
\end{aligned}$$

and, in general,

$$\begin{aligned}
\mathrm{E}(S_R^r) &= \prod_{k=1}^{r}(p+2k-2) \\
&\quad + \frac{1}{24n}\sum_{j=1}^{r}\frac{r!}{(r-j)!}\left[\prod_{k=j+1}^{r}(p+2k-2)\right]2^jA_j. && (2.135)
\end{aligned}$$

Now, from equation (2.99), we can see that

$$\mathrm{E}(S_R) = \mathrm{E}\left[S_0(\mathbf{Z}) + n^{-\frac{1}{2}}S_1(\mathbf{Z}) + n^{-1}S_2(\mathbf{Z})\right] + O\left(n^{-\frac{3}{2}}\right). \tag{2.136}$$

We have

$$
\begin{aligned}
\mathrm{E}(Z_r Z_s) &= \nu_{r,s}, \\
\mathrm{E}(Z_r Z_s Z_t) &= n^{-\frac{1}{2}}\nu_{r,s,t}, \\
\mathrm{E}(Z_r Z_s Z_t Z_u) &= n^{-1}\nu_{r,s,t,u} + \nu_{r,s}\nu_{t,u}[3], \\
\mathrm{E}(Z_r Z_s Z_t Z_u Z_v) &= n^{-\frac{3}{2}}\nu_{r,s,t,u,v} + n^{-\frac{1}{2}}\nu_{r,s}\nu_{t,u,v}[10], \\
\mathrm{E}(Z_r Z_s Z_t Z_u Z_v Z_w) &= n^{-2}\nu_{r,s,t,u,v,w} + n^{-1}\nu_{r,s}\nu_{t,u,v,w}[15] + n^{-1}\nu_{r,s,t}\nu_{u,v,w}[10] \\
&\quad +\nu_{r,s}\nu_{t,u}\nu_{v,w}[15],
\end{aligned}
$$

and so on. Note also that

$$
\begin{aligned}
\nu_{h_1\xi_1,h_2\xi_2} &= \mathrm{E}\left[\left(Z_{\tau_1\xi_1} - \nu_{\tau_1\xi_3}b^{\xi_3\xi_4}Z_{\xi_4\xi_1}\right)\left(Z_{\tau_2\xi_2} - \nu_{\tau_2\xi_5}b^{\xi_5\xi_6}Z_{\xi_6\xi_2}\right)\right] \\
&= \mathrm{E}\left[\left(Z_{\tau_1\xi_1} - \nu_{\tau_1,\xi_3}a^{\xi_3\xi_4}Z_{\xi_4\xi_1}\right)\left(Z_{\tau_2\xi_2} - \nu_{\tau_2,\xi_5}a^{\xi_5\xi_6}Z_{\xi_6\xi_2}\right)\right] \\
&= \nu_{\tau_1\xi_1,\tau_2\xi_2} - \nu_{\tau_1,\xi_3}a^{\xi_3\xi_4}\nu_{\xi_4\xi_1,\tau_2\xi_2} \\
&\quad -\nu_{\xi_2,\xi_5}a^{\xi_5\xi_6}\nu_{\xi_6\xi_2,\tau_1\xi_1} + \nu_{\tau_1,\xi_3}a^{\xi_3,\xi_4}\nu_{\tau_2,\xi_5}a^{\xi_5\xi_6}\nu_{\xi_4\xi_1,\xi_6\xi_2},
\end{aligned}
$$

where $a^{\xi_1\xi_2}$ is, as before, the (ξ_1,ξ_2) entry of the matrix A given by equation (2.50). Using the above and noting that $\nu_{h_1,\xi_1}=0$ (see equation (2.94)), we find that

$$
\begin{aligned}
&\mathrm{E}(S_R) \\
=\;& \nu^{h_1,h_2}\nu_{h_1,h_2} \\
&+n^{-\frac{1}{2}}\Big\{-\nu^{h_1,\phi_1}\nu_{\xi_1,\phi_1,\phi_2}b^{\xi_1\xi_2}\nu^{\phi_2,h_2}n^{-\frac{1}{2}}\nu_{\xi_1,h_1,h_2} \\
&\qquad - 2\nu^{h_1,h_2}b^{\xi_1\xi_2}n^{-\frac{1}{2}}\nu_{\xi_1 h_1,h_2,\xi_2} \\
&\qquad +\nu^{h_1,h_2}\nu_{h_1\xi_2\xi_2}b^{\xi_1\xi_3}b^{\xi_2\xi_4}n^{-\frac{1}{2}}\nu_{\xi_3,\xi_4,h_2}\Big\} \\
&+n^{-1}\Big\{\frac{1}{2}\nu^{h_1,\phi_1}\nu_{\xi_1,\phi_1,\phi_2}b^{\xi_1\xi_2}\nu_{\xi_2\xi_3\xi_4}b^{\xi_3\xi_5}b^{\xi_4\xi_6}\nu^{\phi_2,h_2}\nu_{\xi_5,\xi_6}\nu_{h_1,h_2} \\
&\qquad + \frac{1}{2}\nu^{h_1,\phi_1}\left(\nu_{\phi_1,\phi_2,\xi_1\xi_2} + \nu_{\phi_1,\xi_1,\phi_2\xi_2} + \nu_{\phi_2,\xi_1,\phi_1\xi_2} + \nu_{\phi_1,\phi_2,\xi_1,\xi_2}\right) \\
&\qquad\qquad b^{\xi_1\xi_3}b^{\xi_2\xi_4}\nu^{\phi_2,h_2}\nu_{\xi_3,\xi_4}\nu_{h_1,h_2} \\
&\qquad + \nu^{h_1,\phi_1}\nu_{\xi_1,\phi_1,\phi_2}b^{\xi_1\xi_2}\nu^{\phi_2,\phi_3}\nu_{\xi_3,\phi_3,\phi_4}b^{\xi_3\xi_4}\nu^{\phi_4,h_2}\nu_{\xi_2,\xi_4}\nu_{h_1,h_2} \\
&\qquad + \nu^{h_1,h_2}b^{\xi_1\xi_2}b^{\xi_3\xi_4}\nu_{h_1\xi_1,h_2\xi_3}\nu_{\xi_2,\xi_4} \\
&\qquad +\frac{1}{4}\nu^{h_1,h_2}\nu_{h_1\xi_1\xi_2}b^{\xi_1\xi_3}b^{\xi_2\xi_4}\nu_{h_2\xi_5\xi_6}b^{\xi_5\xi_7}b^{\xi_6\xi_8}\left(\nu_{\xi_3,\xi_4}\nu_{\xi_7,\xi_8}[3]\right)\Big\} \\
&+O\left(n^{-\frac{3}{2}}\right) \qquad (2.137) \\
=\;& \nu^{h_1,h_2}\nu_{h_1,h_2}
\end{aligned}
$$

$$
\begin{aligned}
&+n^{-1}\left\{-\nu^{h_1,\phi_1}\nu_{\xi_1,\phi_1,\phi_2}a^{\xi_1\xi_2}\nu^{\phi_2,h_2}\nu_{\xi_1,h_1,h_2}\right.\\
&\quad+2\nu^{h_1,h_2}a^{\xi_1\xi_2}\nu_{\xi_1h_1,h_2,\xi_2}\\
&\quad+\nu^{h_1,h_2}\nu_{h_1\xi_2\xi_2}a^{\xi_1\xi_3}a^{\xi_2\xi_4}\nu_{\xi_3,\xi_4,h_2}\\
&\quad-\frac{1}{2}\nu^{h_1,\phi_1}\nu_{\xi_1,\phi_1,\phi_2}a^{\xi_1\xi_2}\nu_{\xi_2\xi_3\xi_4}a^{\xi_3\xi_5}a^{\xi_4\xi_6}\nu^{\phi_2,h_2}\nu_{\xi_5,\xi_6}\nu_{h_1,h_2}\\
&\quad+\frac{1}{2}\nu^{h_1,\phi_1}\left(\nu_{\phi_1,\phi_2,\xi_1\xi_2}+\nu_{\phi_1,\xi_1,\phi_2\xi_2}+\nu_{\phi_2,\xi_1,\phi_1\xi_2}+\nu_{\phi_1,\phi_2,\xi_1,\xi_2}\right)\\
&\qquad\qquad a^{\xi_1\xi_3}a^{\xi_2\xi_4}\nu^{\phi_2,h_2}\nu_{\xi_3,\xi_4}\nu_{h_1,h_2}\\
&\quad+\nu^{h_1,\phi_1}\nu_{\xi_1,\phi_1,\phi_2}a^{\xi_1\xi_2}\nu^{\phi_2,\phi_3}\nu_{\xi_3,\phi_3,\phi_4}a^{\xi_3\xi_4}\nu^{\phi_4,h_2}\nu_{\xi_2,\xi_4}\nu_{h_1,h_2}\\
&\quad+\nu^{h_1,h_2}a^{\xi_1\xi_2}a^{\xi_3\xi_4}\nu_{h_1\xi_1,h_2\xi_3}\nu_{\xi_2,\xi_4}\\
&\quad\left.+\frac{1}{4}\nu^{h_1,h_2}\nu_{h_1\xi_1\xi_2}a^{\xi_1\xi_3}a^{\xi_2\xi_4}\nu_{h_2\xi_5\xi_6}a^{\xi_5\xi_7}a^{\xi_6\xi_8}\left(\nu_{\xi_3,\xi_4}\nu_{\xi_7,\xi_8}[3]\right)\right\}\\
&+O\left(n^{-\frac{3}{2}}\right) \qquad (2.138)
\end{aligned}
$$

Consider now terms of the form $\nu^{h_1,h_2}N_{h_1h_2I}$, where $N_{h_1h_2I}$ represents any product of the νs involving the subscripts h_1 and h_2, and any other indices. We can see that

$$
\begin{aligned}
\nu^{h_1,h_2}N_{h_1h_2I} &= \nu^{\tau_1,\tau_2}\left(N_{\tau_1\tau_2I}\right.\\
&\quad-\nu_{\tau_1,\xi_1}a^{\xi_1\xi_2}N_{\xi_2\tau_2I}-\nu_{\tau_2,\xi_3}a^{\xi_3\xi_4}N_{\tau_1\xi_4I}\\
&\quad\left.+\nu_{\tau_1,\xi_1}a^{\xi_1\xi_2}\nu_{\tau_2,\xi_3}a^{\xi_3\xi_4}N_{\xi_2\xi_4I}\right)\\
&= m^{\phi_1\phi_2}N_{\phi_1\phi_2I}, \qquad (2.139)
\end{aligned}
$$

where $m^{\phi_1\phi_2}$ is the (ϕ_1,ϕ_2) entry of the matrix M defined by equation (2.51). Similarly,

$$
\begin{aligned}
\nu^{\phi_1,h_1}N_{h_1\phi_1I} &= \nu^{\phi_1,\tau_1}\left(N_{\tau_1\phi_1I}-\nu_{\tau_1,\xi_1}a^{\xi_1\xi_2}N_{\xi_2\phi_1I}\right)\\
&= \nu^{\tau_2,\tau_1}\left(N_{\tau_1\tau_2I}-\nu_{\tau_1,\xi_1}a^{\xi_1\xi_2}N_{\xi_2\tau_2I}\right)\\
&\quad+\nu^{\xi_3,\tau_1}\left(N_{\tau_1\xi_3I}-\nu_{\tau_1,\xi_1}a^{\xi_1\xi_2}N_{\xi_2\xi_3I}\right)\\
&= \nu^{\tau_2,\tau_1}\left(N_{\tau_1\tau_2I}-\nu_{\tau_1,\xi_1}a^{\xi_1\xi_2}N_{\xi_2\tau_2I}\right.\\
&\quad\left.-a^{\xi_3\xi_4}\nu_{\xi_4,\tau_2}N_{\tau_1\xi_3I}+a^{\xi_3\xi_4}\nu_{\xi_4,\tau_2}\nu_{\tau_1,\xi_1}a^{\xi_1\xi_2}N_{\xi_2\xi_3I}\right)\\
&= m^{\phi_1\phi_2}N_{\phi_2\phi_1I}. \qquad (2.140)
\end{aligned}
$$

Terms of the form $a^{\xi_1\xi_2}N_{\xi_1\xi_2I}$ may be written as $a^{\phi_1\phi_2}N_{\phi_1\phi_2I}$, as all the terms of the partitioned matrix A not in the (ξ,ξ) partition are zero. Using these results we can rewrite (2.138) as

$$
\mathrm{E}(S_R) \;=\; p
$$

$$
\begin{aligned}
&+n^{-1}\Big\{-\nu_{\phi_1,\phi_2,\phi_3}\nu_{\phi_4,\phi_5,\phi_6}m^{\phi_1\phi_2}m^{\phi_3\phi_4}a^{\phi_5\phi_6}\\
&\qquad+2\nu_{\phi_1\phi_2,\phi_3\phi_4}m^{\phi_1\phi_3}a^{\phi_2\phi_4}\\
&\qquad+\nu_{\phi_1,\phi_2,\phi_3}\nu_{\phi_4\phi_5\phi_6}m^{\phi_1\phi_4}a^{\phi_2\phi_5}a^{\phi_3\phi_6}\\
&\qquad-\frac{1}{2}\nu_{\phi_1,\phi_2,\phi_3}\nu_{\phi_4\phi_5\phi_6}a^{\phi_1\phi_4}m^{\phi_2\phi_3}a^{\phi_5\phi_6}\\
&\qquad+\frac{1}{2}\left(\nu_{\phi_1,\phi_2,\phi_3\phi_4}+\nu_{\phi_1,\phi_3,\phi_2\phi_4}+\nu_{\phi_2,\phi_3,\phi_1\phi_4}+\nu_{\phi_1,\phi_2,\phi_3,\phi_4}\right)\\
&\qquad\qquad\qquad m^{\phi_1\phi_2}a^{\phi_3\phi_4}\\
&\qquad+\nu_{\phi_1,\phi_2,\phi_3}\nu_{\phi_4,\phi_5,\phi_6}a^{\phi_1\phi_4}m^{\phi_2\phi_5}m^{\phi_3\phi_6}\\
&\qquad+\nu_{\phi_1\phi_2,\phi_3\phi_4}m^{\phi_1\phi_3}a^{\phi_2\phi_4}\\
&\qquad+\frac{1}{4}\nu_{\phi_1\phi_2\phi_3}\nu_{\phi_4\phi_5\phi_6}m^{\phi_1\phi_4}a^{\phi_2\phi_3}a^{\phi_5\phi_6}\\
&\qquad+\frac{1}{2}\nu_{\phi_1\phi_2\phi_3}\nu_{\phi_4\phi_5\phi_6}m^{\phi_1\phi_4}a^{\phi_2\phi_5}a^{\phi_3\phi_6}\Big\}\\
&+O\left(n^{-\frac{3}{2}}\right). \qquad (2.141)
\end{aligned}
$$

Comparing (2.138) with (2.135) where $r=1$, noting that $\nu_{\phi_1\phi_2\phi_3}=-\nu_{\phi_1,\phi_2,\phi_3}$ and noting that, since all of the summations are from 1 to $m=p+q$, i.e. over the full range of ϕ, the expressions may be written in ordinary index notation, we have

$$
\begin{aligned}
A_1 \;=\;& 12\nu_{rs,tu}m^{rt}a^{su}\\
&+3\nu_{r,s,t}\nu_{u,v,w}m^{ru}a^{st}a^{vw}\\
&+6\nu_{r,s,t}\nu_{u,v,w}m^{ru}a^{sv}a^{tw}\\
&+6\left(\nu_{r,s,tu}+\nu_{r,s,t,u}\right)m^{rs}a^{tu}\\
&+36\nu_{r,s,tu}a^{rt}m^{su}\\
&+6\nu_{r,s,t}\nu_{u,v,w}a^{ru}m^{st}a^{vw}. \qquad (2.142)
\end{aligned}
$$

In order to calculate $\mathrm{E}\left(S_R^r\right)$ for $r\geq 2$, we need the following result.

Proposition 2.2 *If the indices $i_1,i_2,\ldots$ take values $1,\ldots,p$ and*

$$
\begin{aligned}
f(r) &= \prod_{k=1}^{r}\nu^{i_{2k-1},i_{2k}}\left(\prod_{j=1}^{r}\nu_{i_{2j-1},i_{2j}}\left[\frac{(2k)!}{2^k k!}\right]\right), &(2.143)\\
g(r) &= \prod_{k=1}^{r}\nu^{i_{2k-1},i_{2k}}\left(\nu_{i_1,i_2,i_3,i_4}\prod_{k=3}^{r}\nu_{i_{2k-1},i_{2k}}\left[\frac{(2r)!}{2^{r-2}4!(r-2)!}\right]\right), &(2.144)\\
h(r) &= \prod_{k=1}^{r}\nu^{i_{2k-1},i_{2k}}\left(\nu_{i_1,i_2,i_3}\nu_{i_4,i_5,i_6}\prod_{k=4}^{r}\nu_{i_{2k-1},i_{2k}}\left[\frac{(2r)!}{2^{r-2}(3!)^2(r-3)!}\right]\right) &(2.145)
\end{aligned}
$$

then

$$
\begin{aligned}
f(r) &= \prod_{k=1}^{r}(p+2k-2)\,, \quad r \geq 1, &(2.146)\\
g(r) &= \nu^{i_1,i_2}\nu^{i_3,i_4}\nu_{i_1,i_2,i_3,i_4}\begin{pmatrix} k \\ 2 \end{pmatrix}\prod_{k=3}^{r}(p+2k-2), \quad r \geq 2, &(2.147)\\
h(r) &= \left(4\nu^{i_1,i_2}\nu^{i_3,i_4}\nu^{i_5,i_6}+6\nu^{i_1,i_4}\nu^{i_2,i_5}\nu^{i_3,i_6}\right)\nu_{i_1,i_2,i_3}\nu_{i_4,i_5,i_6}\\
&\qquad \begin{pmatrix} k \\ 3 \end{pmatrix}\prod_{k=4}^{r}(p+2k-2), \quad r \geq 3. &(2.148)
\end{aligned}
$$

Proof : Beginning with $f(r)$, we have

$$
\begin{aligned}
f(1) &= \nu^{i_1,i_2}\nu_{i_1,i_2}\\
&= p &(2.149)\\
f(r) &= \nu^{i_1,i_2}\ldots\nu^{i_{2r-3},i_{2r-2}}\nu_{i_1,i_2}\ldots\nu_{i_{2r-3},i_{2r-2}}\nu^{i_{2r-1},i_{2r}}\nu_{i_{2r-1},i_{2r}}\\
&\quad +\nu^{i_1,i_2}\ldots\nu^{i_{2r-3},i_{2r-2}}\left(\nu_{i_1,i_2}\ldots\nu_{i_{2r-3},i_{2r}}\left[\frac{(2r-2)!}{2^{r-1}(r-1)!}\right]\right)\\
&\qquad \nu^{i_{2r-1},i_{2r}}\nu_{i_{2r-1},i_{2r-2}}\left\{\frac{(2r-1)(2r)}{2r}-1\right\}\\
&= f(r-1)(p+2r-2), &(2.150)
\end{aligned}
$$

therefore

$$
f(r)=\prod_{k=1}^{r}(p+2k-2)\,, \quad r \geq 1. \qquad (2.151)
$$

Similarly,

$$
\begin{aligned}
g(2) &= \nu^{i_1,i_2}\nu^{i_3,i_4}\nu_{i_1,i_2,i_3,i_4} &(2.152)\\
g(r) &= \nu^{i_1,i_2}\ldots\nu^{i_{2r-3},i_{2r-2}}\left(\nu_{i_1,i_2,i_3,i_4}\nu_{i_5,i_6}\ldots\nu_{i_{2r-3},i_{2r-2}}\left[\frac{(2r-2)!}{2^{r-3}4!(r-3)!}\right]\right)\\
&\qquad \nu^{i_{2r-1},i_{2r}}\nu_{i_{2r-1},i_{2r}}\frac{r}{r-2}\\
&\quad +\nu^{i_1,i_2}\ldots\nu^{i_{2r-3},i_{2r-2}}\left(\nu_{i_1,i_2,i_3,i_4}\nu_{i_5,i_6}\ldots\nu_{i_{2r-3},i_{2r}}\left[\frac{(2r-2)!}{2^{r-3}4!(r-3)!}\right]\right)\\
&\qquad \nu^{i_{2r-1},i_{2r}}\nu_{i_{2r-2},i_{2r-1}}\left\{\frac{(2r-1)(2r)}{2(r-2)}-\frac{r}{r-2}\right\}\\
&= g(r-1)\frac{r}{r-2}(p+2r-2)\,, &(2.153)
\end{aligned}
$$

so

$$
g(r)=\nu^{i_1,i_2}\nu_{i_3,i_4}\nu_{i_1,i_2,i_3,i_4}\frac{r(r-1)}{2}\prod_{k=3}^{r}(p+2k-2), \quad r \geq 2, \qquad (2.154)
$$

and

$$
\begin{aligned}
h(3) &= \nu^{i_1,i_2}\nu^{i_3,i_4}\nu^{i_5,i_6}\left(\nu_{i_1,i_2,i_3}\nu_{i_4,i_5,i_6}[10]\right) \\
&= \left(4\nu^{i_1,i_2}\nu^{i_3,i_4}\nu^{i_5,i_6} + g\nu^{i_1,i_4}\nu^{i_2,i_5}\nu^{i_3,i_6}\right)\nu_{i_1,i_2,i_3}\nu_{i_4,i_5,i_6} \qquad (2.155) \\
h(r) &= \nu^{i_1,i_2}\ldots\nu^{i_{2r-1},i_{2r}}\left(\nu_{i_1,i_2,i_3}\nu_{i_4,i_5,i_6}\nu_{i_7,i_8}\ldots\nu_{i_{2r-3},i_{2r-2}}\left[\frac{(2r-2)!}{2^{r-3}(3!)^2(r-4)!}\right]\right) \\
&\qquad \nu^{i_{2r-1},i_{2r}}\nu_{i_{2r-1},i_{2r}}\frac{r}{r-3} \\
&\quad +\nu^{i_1,i_2}\ldots\nu^{i_{2r-1},i_{2r}}\left(\nu_{i_1,i_2,i_3}\nu_{i_4,i_5,i_6}\nu_{i_7,i_8}\ldots\nu_{i_{2r-3},i_{2r}}\left[\frac{(2r-2)!}{2^{r-3}(3!)^2(r-4)!}\right]\right) \\
&\qquad \nu^{i_{2r-1},i_{2r}}\nu_{i_{2r-1},i_{2r-2}}\left\{\frac{2r(2r-1)}{2(r-3)} - \frac{r}{r-3}\right\} \\
&= h(r-1)\frac{r}{r-3}(p+2r-2) \qquad (2.156)
\end{aligned}
$$

so

$$
\begin{aligned}
h(r) &= \left(4\nu^{i_1,i_2}\nu^{i_3,i_4}\nu^{i_5,i_6} + 6\nu^{i_1,i_4}\nu^{i_2,i_5}\nu^{i_3,i_6}\right)\nu_{i_1,i_2,i_3}\nu_{i_4,i_5,i_6} \\
&\qquad \binom{k}{3}\prod_{k=4}^{r}(p+2k-2), \quad r \geq 3. \qquad (2.157)
\end{aligned}
$$

◇

From Proposition 2.2 we find that

$$
\begin{aligned}
\mathrm{E}\left(S_0^r\right) &= \nu^{h_1,h_2}\ldots\nu^{h_{2r-1},h_{2r}}\mathrm{E}\left(Z_{h_1}\ldots Z_{h_{2r}}\right) \\
&= \nu^{h_1,h_2}\ldots\nu^{h_{2r-1},h_{2r}} \\
&\quad \left\{\left(\nu_{h_1,h_2}\ldots\nu_{h_{2r-1},h_{2r}}\left[\frac{(2r)!}{2^r r!}\right]\right)\right. \\
&\qquad + n^{-1}\left[\left(\nu_{h_1,h_2,h_3,h_4}\nu_{h_5,h_6}\ldots\nu_{h_{2r-1},h_{2r}}\left[\frac{(2r)!}{2^{r-2}4!(r-2)!}\right]\right)\right. \\
&\qquad\qquad + \left(\nu_{h_1,h_2,h_3}\nu_{h_4,h_5,h_6}\nu_{h_7,h_8}\ldots\nu_{h_{2r-1},h_{2r}}\right. \\
&\qquad\qquad\qquad \left.\left.\left.\left[\frac{(2r)!}{2^{r-2}(3!)^2(r-3)!}\right]\right)\right]\right\} \\
&\quad +O\left(n^{-\frac{3}{2}}\right) \\
&= \prod_{k=1}^{r}(p+2k-2) \\
&\quad +n^{-1}\left\{\nu^{h_1,h_2}\nu^{h_3,h_4}\nu_{h_1,h_2,h_3,h_4}\frac{r(r-1)}{2}\prod_{k=3}^{r}(p+2k-2)\right. \\
&\qquad + \left(4\nu^{h_1,h_2}\nu^{h_3,h_4}\nu^{h_5,h_6} + 6\nu^{h_1,h_4}\nu^{h_2,h_5}\nu^{h_3,h_6}\right)
\end{aligned}
$$

$$\nu_{h_1,h_2,h_3}\nu_{h_4,h_5,h_6}\frac{r(r-1)(r-2)}{6}\prod_{k=4}^{r}(p+2k-2)\Bigg\}$$
$$+O\left(n^{-\frac{3}{2}}\right), \quad r \geq 3, \tag{2.158}$$

and

$$\begin{aligned}
\mathrm{E}\left(S_0^2\right) &= p(p+2) \\
&\quad + n^{-1}\nu^{h_1,h_2}\nu^{h_3,h_4}\nu_{h_1,h_2,h_3,h_4} \\
&\quad + O\left(n^{-\frac{3}{2}}\right).
\end{aligned} \tag{2.159}$$

Also, by analogy with the expressions in Proposition 2.2 it follows that

$$\begin{aligned}
\mathrm{E}\left(S_0^{r-1}S_1\right) &= n^{-\frac{1}{2}}\Bigg\{\nu^{h_1,\phi_1}\nu_{\xi_1,\phi_1,\phi_2}b^{\xi_1\xi_2}\nu^{\phi_2,h}\nu^{h_3,h_4}\ldots\nu^{h_{2r-1},h_{2r}} \\
&\qquad \left(\nu_{\xi_2,h_1,h_2}\nu_{h_3,h_4}\ldots\nu_{h_{2r-1},h_{2r}}\left[\frac{(2r)!}{2^r(r-1)!}\right]\right) \\
&\quad - 2\nu^{h_1,h_2}b^{\xi_1\xi_2}\nu^{h_3,h_4}\ldots\nu^{h_{2r-1},h_{2r}} \\
&\qquad \left(\nu_{h_1\xi_1,\xi_2,h_2}\nu_{h_3,h_4}\ldots\nu_{h_{2r-1},h_{2r}}\left[\frac{(2r-1)!}{2^{r-1}(r-1)!}\right]\right) \\
&\quad + \nu^{h_1,h_2}b^{\xi_1\xi_3}b^{\xi_2\xi_4}\nu_{h_1\xi_1\xi_2}\nu^{h_3,h_4}\ldots\nu^{h_{2r-1},h_{2r}} \\
&\qquad \left(\nu_{\xi_3,\xi_4,h_2}\nu_{h_3,h_4}\ldots\nu_{h_{2r-1},h_{2r}}\left[\frac{(2r-1)!}{2^{r-1}(r-1)!}\right]\right) \\
&\quad + \nu^{h_1,h_2}b^{\xi_1\xi_3}b^{\xi_2\xi_4}\nu_{h_1\xi_1\xi_2}\nu^{h_3,h_4}\ldots\nu^{h_{2r-1},h_{2r}}\nu_{\xi_3,\xi_4} \\
&\qquad \left(\nu_{h_2,h_3,h_4}\nu_{h_5,h_6}\ldots\nu_{h_{2r-1},h_{2r}}\left[\frac{(2r-1)!}{3!\,2^{r-2}(r-2)!}\right]\right)\Bigg\} \\
&\quad + O\left(n^{-\frac{3}{2}}\right) \\
&= n^{-\frac{1}{2}}\Bigg\{\nu^{h_1,\phi_1}\nu_{\xi_1,\phi_1,\phi_2}b^{\xi_1\xi_2}\nu^{\phi_2,h_2}\nu_{\xi_2,h_1,h_2}\prod_{k=3}^{r+1}(p+2k-2) \\
&\quad + \nu^{h_1,\phi_1}\nu_{\xi_1,\phi_1,\phi_2}b^{\xi_1\xi_2}\nu^{\phi_2,h_2}\nu^{h_3,h_4}\nu_{\xi_2,h_3,h_4}\nu_{h_1,h_2} \\
&\qquad (r-1)\prod_{k=3}^{r}(p+2k-2) \\
&\quad - 2\nu^{h_1,h_2}b^{\xi_1\xi_2}\nu_{h_1\xi_1,\xi_2,h_2}\prod_{k=2}^{r}(p+2k-2) \\
&\quad + \nu^{h_1,h_2}b^{\xi_1\xi_3}b^{\xi_2\xi_4}\nu_{h_1\xi_1\xi_2}\nu_{\xi_3,\xi_4,h_2}\prod_{k=2}^{r}(p+2k-2) \\
&\quad + \nu^{h_1,h_2}b^{\xi_1\xi_3}b^{\xi_2\xi_4}\nu_{h_1\xi_1\xi_2}\nu_{\xi_3,\xi_4}\nu^{h_3,h_4}\nu_{h_2,h_3,h_4} \\
&\qquad (r-1)\prod_{k=3}^{r}(p+2k-2)\Bigg\} \\
&\quad + O\left(n^{-\frac{3}{2}}\right), \quad r \geq 2
\end{aligned} \tag{2.160}$$

$$
\begin{aligned}
\mathrm{E}\left(S_0^{r-1}S_2\right) &= \frac{1}{2}\nu^{h_1,\phi_1}\nu_{\xi_1,\phi_1,\phi_2}b^{\xi_1\xi_2}\nu_{\xi_2\xi_3\xi_4}b^{\xi_3\xi_5}b^{\xi_4\xi_6}\nu^{\phi_2,h_2}\nu^{h_3,h_4}\ldots\nu^{h_{2r-1},h_{2r}}\nu_{\xi_5,\xi_6} \\
&\qquad \left(\nu_{h_1,h_2}\ldots\nu_{h_{2r-1},h_{2r}}\left[\frac{(2r)!}{2^r r!}\right]\right) \\
&\quad +\frac{1}{2}\nu^{h_1,\phi_1}\left(\nu_{\phi_1,\phi_2,\xi_1\xi_2}+2\nu_{\phi_1,\xi_1,\phi_2\xi_2}+\nu_{\phi_1,\phi_2,\xi_1,\xi_2}\right)b^{\xi_1\xi_3}b^{\xi_2\xi_4}\nu^{\phi_2,h_2} \\
&\qquad \nu^{h_3,h_4}\ldots\nu^{h_{2r-1},h_{2r}}\nu_{\xi_3,\xi_4}\left(\nu_{h_1,h_2}\ldots\nu_{h_{2r-1},h_{2r}}\left[\frac{(2r)!}{2^r r!}\right]\right) \\
&\quad +\nu^{h_1,\phi_1}\nu_{\xi_1,\phi_1,\phi_2}b^{\xi_1\xi_2}\nu^{\phi_2,\phi_3}\nu_{\xi_3,\phi_3,\phi_4}b^{\xi_3\xi_4}\nu^{\phi_4,h_2}\nu^{h_3,h_4}\ldots\nu^{h_{2r-1},h_{2r}}\nu_{\xi_2,\xi_4} \\
&\quad \left(\nu_{h_1,h_2}\ldots\nu_{h_{2r-1},h_{2r}}\left[\frac{(2r)!}{2^r r!}\right]\right) \\
&\quad +\nu^{h_1,h_2}b^{\xi_1\xi_2}b^{\xi_3\xi_4}\nu^{h_3,h_4}\ldots\nu^{h_{2r-1},h_{2r}}\nu_{h_1\xi_1,h_2\xi_3}\nu_{\xi_2,\xi_4} \\
&\quad \left(\nu_{h_3,h_4}\ldots\nu_{h_{2r-1},h_{2r}}\left[\frac{(2r-2)!}{2^{r-1}(r-1)!}\right]\right) \\
&\quad +\frac{1}{4}\nu^{h_1,h_2}\nu_{h_1\xi_1\xi_2}b^{\xi_1\xi_3}b^{\xi_2\xi_4}\nu_{h_2\xi_5\xi_6}b^{\xi_5\xi_7}b^{\xi_6\xi_8}\nu^{h_3,h_4}\ldots\nu^{h_{2r-1},h_{2r}} \\
&\qquad \left(\nu_{\xi_3,\xi_4}\nu_{\xi_7,\xi_8}[3]\right)\left(\nu_{h_3,h_4}\ldots\nu_{h_{2r-1},h_{2r}}\left[\frac{(2r-2)!}{2^{r-1}(r-1)!}\right]\right) \\
&\quad +O\left(n^{-1}\right) \\
&= \frac{1}{2}\nu^{h_1,\phi_1}\nu_{\xi_1,\phi_1,\phi_2}b^{\xi_1\xi_2}\nu_{\xi_2\xi_3\xi_4}b^{\xi_3\xi_5}b^{\xi_4\xi_6}\nu^{\phi_2,h_2}\nu_{\xi_5,\xi_6}\nu_{h_1,h_2}\prod_{k=2}^{r}(p+2k-2) \\
&\quad +\frac{1}{2}\nu^{h_1,\phi_1}\left(\nu_{\phi_1,\phi_2,\xi_1\xi_2}+2\nu_{\phi_1,\xi_1,\phi_2\xi_2}+\nu_{\phi_1,\phi_2,\xi_1,\xi_2}\right) \\
&\qquad b^{\xi_1\xi_3}b^{\xi_2\xi_4}\nu^{\phi_2,h_2}\nu_{\xi_3,\xi_4}\nu_{h_1,h_2} \\
&\qquad\qquad \prod_{k=2}^{r}(p+2k-2) \\
&\quad +\nu^{h_1,\phi_1}\nu_{\xi_1,\phi_1,\phi_2}b^{\xi_1\xi_2}\nu^{\phi_2,\phi_3}\nu_{\xi_3,\phi_3,\phi_4}b^{\xi_3\xi_4}\nu^{\phi_4,h_2}\nu_{\xi_2,\xi_4}\nu_{h_1,h_2} \\
&\qquad\qquad \prod_{k=2}^{r}(p+2k-2) \\
&\quad +\nu^{h_1,h_2}b^{\xi_1\xi_2}b^{\xi_3\xi_4}\nu_{h_1\xi_1,h_2\xi_3}\nu_{\xi_2,\xi_4}\prod_{k=1}^{r-1}(p+2k-2) \\
&\quad +\frac{1}{4}\nu^{h_1,h_2}\nu_{h_1\xi_1\xi_2}b^{\xi_1\xi_3}b^{\xi_2\xi_4}\nu_{h_2\xi_5\xi_6}b^{\xi_5\xi_7}b^{\xi_6\xi_8}\left(\nu_{\xi_3,\xi_4}\nu_{\xi_7,\xi_8}[3]\right) \\
&\qquad\qquad \prod_{k=1}^{r-1}(p+2k-2) \\
&\quad +O\left(n^{-1}\right),\quad r\geq 2 \qquad (2.161)
\end{aligned}
$$

$$
\begin{aligned}
\mathrm{E}\left(S_0^{r-2}S_1^2\right) &= \nu^{h_1,\phi_1}\nu_{\xi_1,\phi_1,\phi_2}b^{\xi_1,\xi_2}\nu^{\phi_2,h_2}\nu^{h_3,\phi_3}\nu_{\xi_3,\phi_3,\phi_4}b^{\xi_3\xi_4}\nu^{\phi_4,h_4} \\
&\qquad \nu^{h_5,h_6}\ldots\nu^{h_{2r-1},h_{2r}}\nu_{\xi_2,\xi_4}\left(\nu_{h_1,h_2}\ldots\nu_{h_{2r-1},h_{2r}}\left[\frac{(2r)!}{2^r r!}\right]\right) \\
&\quad +4\nu^{h_1,h_2}b^{\xi_1\xi_2}\nu^{h_3,h_4}b^{\xi_3\xi_4}\nu^{h_5,h_6}\ldots\nu^{h_{2r-1},h_{2r}}\nu_{h_1\xi_1,h_3\xi_3}\nu_{\xi_2,\xi_4}
\end{aligned}
$$

$$
\begin{aligned}
& \left(\nu_{h_2,h_4}\nu_{h_5,h_6}\ldots\nu_{h_{2r-1},h_{2r}}\left[\frac{(2r-2)!}{2^{r-1}(r-1)!}\right]\right) \\
& +\nu^{h_1,h_2}b^{\xi_1\xi_3}b^{\xi_2\xi_4}\nu_{h_1\xi_1\xi_2}\nu^{h_3,h_4}b^{\xi_5\xi_7}b^{\xi_6\xi_8}\nu_{h_3\xi_5\xi_6}\nu^{h_5,h_6}\ldots\nu^{h_{2r-1},h_{2r}} \\
& \quad (\nu_{\xi_3,\xi_4}\nu_{\xi_7,\xi_8}[3])\left(\nu_{h_2,h_4}\nu_{h_5,h_6}\ldots\nu_{h_{2r-1},h_{2r}}\left[\frac{(2r-2)!}{2^{r-1}(r-1)!}\right]\right) \\
= \; & \nu^{h_1,\phi_1}\nu_{\xi_1,\phi_1,\phi_2}b^{\xi_1\xi_2}\nu^{\phi_2,h_2}\nu^{h_3,\phi_3}\nu_{\xi_3,\phi_3,\phi_4}b^{\xi_3\xi_4}\nu^{\phi_4,h_4}\nu_{\xi_2,\xi_4} \\
& \qquad (\nu_{h_1,h_2}\nu_{h_3,h_4}[3])\prod_{k=3}^{r}(p+2k-2) \\
& +4\nu^{h_1,h_2}b^{\xi_1\xi_2}\nu^{h_3,h_4}b^{\xi_3\xi_4}\nu_{h_1\xi_1,h_3\xi_3}\nu_{\xi_2,\xi_4}\nu_{h_2,h_4}\prod_{k=2}^{r-1}(p+2k-2) \\
& +\nu^{h_1,h_3}b^{\xi_1\xi_3}b^{\xi_2\xi_4}\nu_{h_1\xi_1\xi_2}b^{\xi_5\xi_7}b^{\xi_6\xi_8}\nu_{h_3\xi_5\xi_6} \\
& \quad (\nu_{\xi_3,\xi_4}\nu_{\xi_7,\xi_8}[3])\prod_{k=2}^{r-1}(p+2k-2) \\
& +O\left(n^{-1}\right), \quad r\geq 2.
\end{aligned}
\tag{2.162}
$$

Now, for $r \geq 2$,

$$
\begin{aligned}
\mathrm{E}\left(S_R^r\right) = \; & \mathrm{E}\left(S_0^r\right) + n^{-\frac{1}{2}} r \mathrm{E}\left(S_0^{r-1}S_1\right) \\
& +n^{-1}\left\{r\mathrm{E}\left(S_0^{r-1}S_2\right) + r(r-1)\mathrm{E}\left(S_0^{r-2}S_1^2\right)\right\} \\
& +O\left(n^{-\frac{3}{2}}\right).
\end{aligned}
\tag{2.163}
$$

Substituting (2.158) to (2.162) into (2.163) we have, for $r \geq 3$,

$$
\begin{aligned}
\mathrm{E}\left(S_R^r\right) = \; & \prod_{k=1}^{r}(p+2k-2) \\
& +n^{-1}\left\{\frac{1}{2}\nu^{h_1,h_2}\nu^{h_3,h_4}\nu_{h_1,h_2,h_3,h_4}r(r-1)\prod_{k=3}^{r}(p+2k-2)\right. \\
& \quad +\left(\frac{2}{3}\nu^{h_1,h_2}\nu^{h_3,h_4}\nu^{h_5,h_6}+\nu^{h_1,h_4}\nu^{h_2,h_5}\nu^{h_3,h_6}\right)\nu_{h_1,h_2,h_3}\nu_{h_4,h_5,h_6} \\
& \qquad r(r-1)(r-2)\prod_{k=4}^{r}(p+2k-2) \\
& \quad +\frac{1}{2}\nu^{h_1,\phi_1}\left(\nu_{\phi_1,\phi_2,\xi_1\xi_2}+2\nu_{\phi_1,\xi_1,\phi_2\xi_2}+\nu_{\phi_1,\phi_2,\xi_1,\xi_2}\right)b^{\xi_1\xi_3}b^{\xi_2\xi_4} \\
& \qquad \nu^{\phi_2,h_2}\nu_{\xi_3,\xi_4}\nu_{h_1,h_2}r\prod_{k=2}^{r}(p+2k-2) \\
& \quad +\nu^{h_1,\phi_1}\nu_{\xi_1,\phi_1,\xi_5}b^{\xi_1\xi_2}a^{\xi_5,\xi_6}\nu_{\xi_3,\xi_6,\phi_4}b^{\xi_3\xi_4}\nu^{\phi_4,h_2}\nu_{\xi_2,\xi_4}\nu_{h_1,h_2} \\
& \qquad r\prod_{k=2}^{r}(p+2k-2) \\
& \quad +\nu^{h_1,h_2}b^{\xi_1\xi_2}b^{\xi_3\xi_4}\nu_{h_1\xi_1,h_2\xi_3}\nu_{\xi_2,\xi_4}r\prod_{k=2}^{r}(p+2k-2)
\end{aligned}
$$

$$
\begin{aligned}
&+\frac{1}{4}\nu^{h_1,h_2}\nu_{h_1\xi_1\xi_2}b^{\xi_1\xi_3}b^{\xi_2\xi_4}\nu_{h_2\xi_5\xi_6}b^{\xi_5\xi_7}b^{\xi_6\xi_8}\left(\nu_{\xi_3,\xi_4}\nu_{\xi_7,\xi_8}[3]\right)\\
&\qquad r\prod_{k=2}^{r}(p+2k-2)\\
&+\frac{1}{2}\nu^{h_1,\phi_1}\nu_{\xi_1,\phi_1,\phi_2}b^{\xi_1\xi_2}b^{\xi_1\xi_2}\nu^{\phi_2,h_2}\nu^{h_3,\phi_3}\nu_{\xi_3,\phi_3,\phi_4}b^{\xi_3\xi_4}\nu^{\phi_4,h_4}\nu_{\xi_2,\xi_4}\\
&\qquad\left(\nu_{h_1,h_2}\nu_{h_3,h_4}[3]\right)r(r-1)\prod_{k=3}^{r}(p+2k-2)\\
&+2\nu^{h_1,\phi_1}\nu_{\xi_1,\phi_1,\phi_2}b^{\xi_1\xi_2}\nu^{\phi_2,h_2}\nu_{\xi_2,h_1,h_2}r(r-1)\prod_{k=3}^{r}(p+2k-2)\\
&+\nu^{h_1,\phi_1}\nu_{\xi_1,\phi_1,\phi_2}b^{\xi_1\xi_2}\nu^{\phi_2,h_2}\nu^{h_3,h_4}\nu_{\xi_2,h_3,h_4}\nu_{h_1,h_2}\\
&\qquad r(r-1)\prod_{k=3}^{r}(p+2k-2)\\
&-2\nu^{h_1,h_2}b^{\xi_1\xi_2}\nu_{h_1\xi_1,\xi_2,h_2}r\prod_{k=2}^{r}(p+2k-2)\\
&+\nu^{h_1,h_2}b^{\xi_1\xi_3}b^{\xi_2\xi_4}\nu_{h_1\xi_1\xi_2}\nu_{\xi_3,\xi_4,h_2}r\prod_{k=2}^{r}(p+2k-2)\\
&+\nu^{h_1,h_2}b^{\xi_1\xi_3}b^{\xi_2\xi_4}\nu_{h_1\xi_1\xi_2}\nu_{\xi_3,\xi_4}\nu^{h_3,h_4}\nu_{h_2,h_3,h_4}\\
&\qquad r(r-1)\prod_{k=3}^{r}(p+2k-2)\Bigg\}. \qquad (2.164)
\end{aligned}
$$

From (2.158) to (2.162) and (2.135) we find

$$
\begin{aligned}
A_2 &= 3\nu_{r,s,t,u}m^{rs}m^{tu}\\
&\quad -6\nu_{r,s,t}\nu_{u,v,w}a^{ru}m^{sv}m^{tw}\\
&\quad -3\nu_{r,s,t}\nu_{u,v,w}a^{ru}m^{st}m^{vw}\\
&\quad -6\nu_{r,s,t}\nu_{u,v,w}m^{ru}a^{st}m^{vw}, \qquad (2.165)\\
A_3 &= 3\nu_{r,s,t}\nu_{u,v,w}m^{ru}m^{st}m^{vw}\\
&\quad +2\nu_{r,s,t}\nu_{u,v,w}m^{ru}m^{sv}m^{tw}. \qquad (2.166)
\end{aligned}
$$

Finally, from (2.164), (2.165), (2.166) and (2.135) it follows that

$$
A_r = 0 \quad \text{for } r \geq 4. \qquad (2.167)
$$

Rewriting (2.142), (2.165) and (2.166) in a general parameterization, we have

$$
\begin{aligned}
A_1 &= 12\left(\kappa_{ij,kl}-\kappa^{m,n}\kappa_{m,ij}\kappa_{n,kl}\right)m^{ik}a^{jl}\\
&\quad +3\kappa_{i,j,k}\kappa_{l,m,n}m^{il}a^{jk}a^{mn}\\
&\quad +6\kappa_{i,j,k}\kappa_{l,m,n}m^{il}a^{jm}a^{kn}
\end{aligned}
$$

$$
\begin{aligned}
& +6\left(\kappa_{i,j,kl}-\kappa^{m,n}\kappa_{n,kl}\kappa_{m,i,j}\right)m^{ij}a^{kl} \\
& +6\kappa_{i,j,k,l}m^{ij}a^{kl} \\
& +36\left(\kappa_{i,j,kl}-\kappa^{m,n}\kappa_{n,kl}\kappa_{m,i,j}\right)m^{ik}a^{jl} \\
& +6\kappa_{i,j,k}\kappa_{l,m,n}a^{il}m^{jk}a^{mn} \qquad (2.168) \\
A_2 \;=\; & 3\kappa_{i,j,k,l}m^{ij}m^{kl} \\
& -6\kappa_{i,j,k}\kappa_{l,m,n}a^{il}m^{jm}m^{kn} \\
& -3\kappa_{i,j,k}\kappa_{l,m,n}a^{il}m^{jk}m^{mn} \\
& -6\kappa_{i,j,k}\kappa_{l,m,n}m^{il}a^{jk}m^{mn} \qquad (2.169) \\
A_3 \;=\; & 3\kappa_{i,j,k}\kappa_{l,m,n}m^{il}m^{jk}m^{mn} \\
& +2\kappa_{i,j,k}\kappa_{l,m,n}m^{il}m^{jm}m^{kn}. \qquad (2.170)
\end{aligned}
$$

These expressions can be seen to be invariant under reparameterization.

2.2.8 Examples of Bartlett Adjustment

Some examples of generalized Bartlett adjustments are given by Cordeiro & Ferrari (1991), using the expressions for A_1,A_2 and A_3 derived by Harris (1985). As we have shown here that the expressions for A_1 and A_2 given by Harris are incorrect, it is of interest to reconsider these examples.

Cordeiro and Ferrari discuss two special cases: a one-parameter model, and a model with two orthogonal parameters. For the one-parameter model, under a simple null hypothesis, they note that $A_1 = 0$ and derive a simple expression for the improved score statistic in terms of skewness and kurtosis of the total score function. This expression is correct, as the expression given by Harris for A_2 and that given by equation (2.169) are equal in this case. In fact this is true for any model under a simple null hypothesis. To see this, note that the matrix A is zero in the case of a simple null hypothesis as it depends only on the cumulants of derivatives with respect to the nuisance parameters. From this it clearly follows that $A_1 = 0$, while equation (2.169) for A_2 and equation (2.48), which gives the expression for A_2 derived by Harris, both reduce to

$$A_2 = 3\kappa_{i,j,k,l}m^{ij}m^{kl}. \qquad (2.171)$$

The second case which Cordeiro and Ferrari consider, that of a model with two orthogonal parameters where one is a nuisance parameter, is more complicated. They derive general expressions for A_1, A_2 and A_3 in this situation, and then apply these to several examples. The details are somewhat complicated, but after some algebra it is possible to show that the expressions given for A_2 and A_3 are correct, but that given for A_1 is incorrect. As a specific example, we can look at the first model in Table 2 of Cordeiro & Ferrari (1991), namely the normal distribution parameterized by mean, μ, and dispersion, ϕ, where we test the null hypothesis $H_0 : \phi = \phi^{(0)}$ with unknown mean. Cordeiro and Ferrari note that, given independent observations $y_1, \ldots, y_n$ from this distribution, the score test statistic is

$$S_R = \frac{1}{2n}\left\{n\ \phi^{(0)} \sum_{i=1}^{n} (y_i - \overline{y})^2\right\}^2, \tag{2.172}$$

and they give the improved score statistic as

$$S_R^* = S_R\left\{1 - \frac{1}{18n}\left(33 - 34S_R + 4S_R^2\right)\right\}. \tag{2.173}$$

Using equations (2.168)-(2.170), we find that the correct expression for the improved score statistic is

$$S_R^* = S_R\left\{1 - \frac{1}{18n}\left(-165 - 34S_R + 4S_R^2\right)\right\}. \tag{2.174}$$

Chapter 3

Generalized Bartlett Adjustment

3.1 Conditions for Generalized Bartlett Adjustment

In this section we derive a necessary and sufficient condition for the existence of a generalized Bartlett adjustment. This will be used in Chapter 4 to obtain the generalized Bartlett adjustment for the score test statistic under a mis-specified model.

Consider a statistic X with unknown probability density function $g(x)$ and moment generating function

$$M_X(t) = M_0(t) + n^{-1}M_1(t) + O\left(n^{-\frac{3}{2}}\right), \tag{3.1}$$

where $M_0(t)$ is the moment generating function of some known distribution with probability density function $f(x)$. We may construct a new statistic, X', by taking

$$X' = \left(1 + \frac{1}{n}\sum_{r=0}^{d} a_r X^r\right) X, \tag{3.2}$$

i.e. we use a multiplicative correction factor which is a polynomial in the statistic itself, as described for the particular case of the score test statistic by Cordeiro & Ferrari (1991). If we can find values of d, the degree of the polynomial, and the coefficients $a_0, \ldots, a_d$ such that the distribution of X' has probability density function $f(x)$ with error of order $O\left(n^{-\frac{3}{2}}\right)$ or smaller then we call the process of obtaining the modified statistic X' from the statistic X generalized Bartlett adjustment.

The modified statistic X' has density function $f(x)$ with error of order $O\left(n^{-\frac{3}{2}}\right)$ or smaller if and only if

$$M_{X'}(t) = M_0(t) + O\left(n^{-\frac{3}{2}}\right). \tag{3.3}$$

Now

$$
\begin{aligned}
M_{X'}(t) &= E_g\left[e^{X't}\right] \\
&= E_g\left[e^{\left(1+\frac{1}{n}\sum_{r=0}^{d} a_r X^r\right)Xt}\right] \\
&= E_g\left[e^{Xt}\left\{1+\frac{1}{n}t\sum_{r=0}^{d} a_r X^{r+1}\right\}\right] + O\left(n^{-2}\right) \\
&= M_0(t) + n^{-1}\left\{E_g\left[te^{Xt}\sum_{r=0}^{d} a_r X^{r+1}\right] + M_1(t)\right\} + O\left(n^{-\frac{3}{2}}\right), \quad (3.4)
\end{aligned}
$$

and

$$
\begin{aligned}
&E_g\left[te^{Xt}\sum_{r=0}^{d} a_r X^{r+1}\right] \\
&= t\sum_{r=0}^{d} a_r E_g\left[e^{Xt}X^{r+1}\right] \\
&= t\sum_{r=0}^{d} a_r \int e^{xt}x^{r+1}g(x)\,dx \\
&= t\sum_{r=0}^{d} a_r \frac{d^{(r+1)}}{dt^{(r+1)}}\left\{\int e^{xt}g(x)\,dx\right\} \\
&= t\sum_{r=0}^{d} a_r \frac{d^{(r+1)}}{dt^{(r+1)}} M_X(t), \quad (3.5)
\end{aligned}
$$

so

$$
M_{X'}(t) = M_0(t) + n^{-1}\left\{t\sum_{r=0}^{d} a_r \frac{d^{(r+1)}}{dt^{(r+1)}} M_0(t) + M_1(t)\right\} + O\left(n^{-\frac{3}{2}}\right). \quad (3.6)
$$

Comparing equation (3.3) with equation (3.6) we can see that X' has probability density function $f(x)$ to the required accuracy if and only if

$$
t\sum_{r=0}^{d} a_r \frac{d^{(r+1)}}{dt^{(r+1)}} M_0(t) + M_1(t) = 0, \quad (3.7)
$$

i.e. if and only if $\frac{M_1(t)}{t}$ may be expressed as a linear combination of the derivatives of $M_0(t)$.

Not only is this a necessary and sufficient condition for the existence of a generalized Bartlett adjustment, it also suggests a possible approach to deriving such an adjustment where it does exist, i.e. via the order $O(1)$ and order $O\left(n^{-\frac{1}{2}}\right)$ terms of the moment generating function. This approach is simpler, in some cases, than the direct method used to obtain the Bartlett adjustment to the score statistic in Chapter 2. It will be used in Chapter 4 to derive a generalized Bartlett adjustment to the score statistic under a mis-specified model.

3.2 Example: Application to Weighted Sum of two Bartlett-Adjustable Statistics

As an example of the application of condition (3.7), consider the problem of whether a Bartlett factor exists for the weighted sum of two statistics, each of which has an asymptotic chi-squared distribution and is itself Bartlett-adjustable. This is a problem of some mathematical interest, but is also relevant to consideration of whether a generalized Bartlett adjustment exists for the likelihood ratio test statistic under a mis-specified model. This will be discussed in Chapter 4.

Let X and Y be statistics, based on independent samples of size n, with asymptotic distributions χ_p^2 and χ_q^2 respectively, and let b_X and b_Y be Bartlett factors such that

$$\begin{aligned} \left(1+\frac{b_X}{n}\right)^{-1} X &\sim \chi_p^2, \\ \left(1+\frac{b_Y}{n}\right)^{-1} Y &\sim \chi_q^2, \end{aligned}$$

with error of order $O\left(n^{-\frac{3}{2}}\right)$. Given real constants λ and μ, we define the statistic Z as

$$Z = \lambda X + \mu Y, \tag{3.8}$$

then ask whether it is possible to find a modified statistic, Z', using generalized Bartlett adjustment, i.e.

$$Z' = \left\{1 + \frac{1}{n}\sum_{r=0}^{d} a_r Z^r\right\} Z, \tag{3.9}$$

such that

$$Z' \sim \lambda U_p + \mu U_q \tag{3.10}$$

with error of order $O\left(n^{-\frac{3}{2}}\right)$, where U_p is a χ_p^2-variate and U_q is a χ_q^2-variate.

We have

$$M_{(1+b_X/n)^{-1}X}(t) = (1-2t)^{-\frac{p}{2}} + O\left(n^{-\frac{3}{2}}\right),$$

so

$$\begin{aligned} M_X(t) &= \left[1 - 2\left(1+\frac{b_X}{n}\right)t\right]^{-\frac{p}{2}} + O\left(n^{-\frac{3}{2}}\right) \\ &= \left[1 - 2t - \frac{2b_X}{n}t\right]^{-\frac{p}{2}} + O\left(n^{-\frac{3}{2}}\right) \end{aligned}$$

$$
\begin{aligned}
&= (1-2t)^{-\frac{p}{2}}\left[1-\frac{2b_X t}{(1-2t)\,n}\right]^{-\frac{p}{2}} + O\left(n^{-\frac{3}{2}}\right) \\
&= (1-2t)^{-\frac{p}{2}}\left[1+\frac{pb_X t}{(1-2t)\,n}\right] + O\left(n^{-\frac{3}{2}}\right) \qquad (3.11)
\end{aligned}
$$

and similarly

$$
M_Y(t) = (1-2t)^{-\frac{q}{2}}\left[1+\frac{qb_Y t}{(1-2t)\,n}\right] + O\left(n^{-\frac{3}{2}}\right). \qquad (3.12)
$$

From (3.11) and (3.12) we have

$$
\begin{aligned}
M_Z(t) &= M_{\lambda X+\mu Y}(t) \\
&= (1-2\lambda t)^{-\frac{p}{2}}(1-2\mu t)^{-\frac{q}{2}}\left[1+\frac{pb_X\lambda t}{(1-2\lambda t)\,n}\right]\left[1+\frac{qb_Y\mu t}{(1-2\mu t)\,n}\right] + O\left(n^{-\frac{3}{2}}\right) \\
&= (1-2\lambda t)^{-\frac{p}{2}}(1-2\mu t)^{-\frac{q}{2}}\left[1+\frac{pb_X\lambda t}{(1-2\lambda t)\,n}+\frac{qb_Y\mu t}{(1-2\mu t)\,n}\right] + O\left(n^{-\frac{3}{2}}\right) \\
&= (1-2\lambda t)^{-\frac{p}{2}}(1-2\mu t)^{-\frac{q}{2}} \\
&\quad +\frac{1}{n}\left[pb_X\lambda t(1-2\lambda t)^{-\frac{p+2}{2}}(1-2\mu t)^{-\frac{q}{2}}\right. \\
&\qquad \left. + qb_Y\mu t(1-2\lambda t)^{-\frac{p}{2}}(1-2\mu t)^{-\frac{q+2}{2}}\right] + O\left(n^{-\frac{3}{2}}\right) \qquad (3.13)
\end{aligned}
$$

so

$$
\begin{aligned}
\frac{M_1(t)}{t} &= pb_X\lambda(1-2\lambda t)^{-\frac{p+2}{2}}(1-2\mu t)^{-\frac{q}{2}} \\
&\quad + qb_Y\mu(1-2\lambda t)^{-\frac{p}{2}}(1-2\mu t)^{-\frac{q+2}{2}} \qquad (3.14) \\
M_0(t) &= (1-2\lambda t)^{-\frac{p}{2}}(1-2\mu t)^{-\frac{q}{2}}. \qquad (3.15)
\end{aligned}
$$

If we differentiate $M_0(t)$ we find that

$$
\begin{aligned}
M_0^{(k)}(t) &= \sum_{j=0}^{k}\binom{k}{j}\left(\prod_{i=0}^{j-1} p+2i\right)\left(\prod_{l=0}^{k-j-1} q+2l\right)\lambda^j\mu^{k-j} \\
&\qquad (1-2\lambda t)^{-\frac{p+2j}{2}}(1-2\mu t)^{-\frac{q+2(k-j)}{2}}. \qquad (3.16)
\end{aligned}
$$

Generalized Bartlett adjustment is possible if and only if the equation (3.7) can be satisfied. In this case, (3.7) may be written in the form

$$
\begin{aligned}
&\sum_{r=0}^{d} a_r \sum_{j=0}^{r+1}\binom{r+1}{j}\left(\prod_{i=0}^{j-1} p+2i\right)\left(\prod_{l=0}^{r-j} q+2l\right)\lambda^j\mu^{r+1-j} \\
&\qquad (1-2\lambda t)^{-\frac{p+2j}{2}}(1-2\mu t)^{-\frac{q+2r-2j+2}{2}} \\
&= -pb_X\lambda(1-2\lambda t)^{-\frac{p+2}{2}}(1-2\mu t)^{-\frac{q}{2}} - qb_Y\mu(1-2\lambda t)^{-\frac{p}{2}}(1-2\mu t)^{-\frac{q+2}{2}}, \qquad (3.17)
\end{aligned}
$$

which may be reduced to

$$\begin{aligned}
&\sum_{r=0}^{d} a_r \sum_{j=0}^{r+1} \binom{r+1}{j} \left(\prod_{i=0}^{j-1} p+2i\right) \left(\prod_{l=0}^{r-j} q+2l\right) \lambda^j \mu^{r+1-j} \\
&\qquad (1-2\lambda t)^{-j}(1-2\mu t)^{-(r-j+1)} \\
= &\ -pb_X\lambda(1-2\lambda t)^{-1} - qb_Y\mu(1-2\mu t)^{-1}. \qquad (3.18)
\end{aligned}$$

Equation (3.18) may be satisfied for all t only under certain conditions. Clearly, we must take $d = 0$, so that (3.18) reduces to

$$\begin{aligned}
&a_0 \left\{ q\mu(1-2\mu t)^{-1} + p\lambda(1-2\lambda t)^{-1} \right\} \\
= &\ -pb_X\lambda(1-2\lambda t)^{-1} - qb_Y\mu(1-2\mu t)^{-1}. \qquad (3.19)
\end{aligned}$$

We can find a value of a_0 to satisfy (3.19) for all t if and only if at least one of the following three conditions holds

$$\mu = \lambda \qquad (3.20)$$

$$b_X = b_Y. \qquad (3.21)$$

(3.21) is the trivial case where the Bartlett factors for statistics X and Y are equal. (3.20) shows that Bartlett adjustment of the (unweighted) sum of statistics X and Y, and indeed of any multiple of it, is possible. In all of these cases, we are using "ordinary", i.e. linear, Bartlett adjustment, as the degree of the polynomial in the generalized Bartlett factor is zero.

Chapter 4

Bartlett Adjustment under a Mis-specified Model

Consider the situation, discussed in Chapter 2, where we have independent, identically distributed observations $X_1, \ldots, X_n$ with underlying density $g(x)$, which we summarize by fitting a parametric model $\mathcal{F} = \{f(x;\theta) : \theta \in \Theta\}$. To test the null hypothesis $H_0 : \theta \in \Theta_0$ against the alternative hypothesis $H_1 : \theta \in \Theta_1$, where $\Theta_0 \subset \Theta_1$, we may calculate the likelihood ratio test statistic or the score test statistic, which are, under the null hypothesis, asymptotically distributed as χ^2_d, where $d = \dim \Theta_1 - \dim \Theta_0$.

Now consider the situation where we cannot assume that the null hypothesis is true, i.e. where we cannot assume that the true unknown density $g(x)$ is a member of the parametric family $\mathcal{F}$. Distributional results derived without this assumption can be said to hold for a *mis-specified model.* Ordinarily, in this situation, we might use non-parametric methods to make inference about the data, but we suppose here that the parameters of the model $\mathcal{F}$ are a convenient means of summarizing the data that we wish to keep.

It is clearly necessary to make some assumptions about the true, unknown density $g(x)$, even if we do not assume that it belongs to the parametric family $\mathcal{F}$. The results here follow on from those given by Kent (1982), where the assumption is that the density in $\mathcal{F}$ "nearest" to the density g satisfies the null hypothesis. More rigorously, Kent introduces the *Fraser information*, motivated by Fraser(1965) and defined as

$$F(\theta) = \int \log f(x;\theta) g(x)\, dx. \tag{4.1}$$

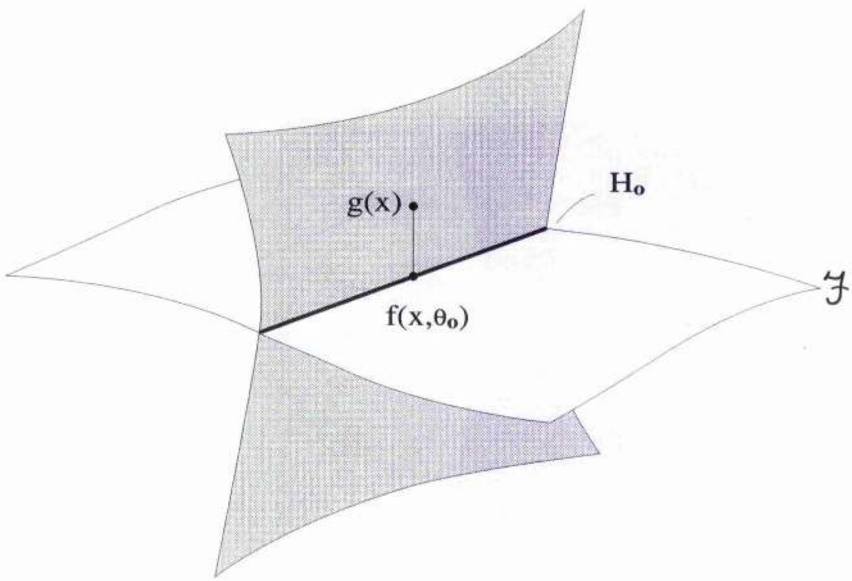

Figure 4.1: Schematic representation of mis-specified model

The value $\theta(g)$ is then defined by

$$F\{\theta(g)\} = \max\{F(\theta) : \theta \in \Theta\}, \tag{4.2}$$

so that $\theta(g)$ is the theoretical analogue of the maximum likelihood estimator for θ. It is assumed that the true density satisfies

$$\theta(g) \in \Theta_0. \tag{4.3}$$

In practice, this assumption reduces to requiring that the assumed model has some features in common with the underlying distribution. This is reasonable, as these are likely to be features observed in the data – for example, one would be unlikely to choose a family of bimodal distributions to model a data set which clearly contained observations from a unimodal distribution.

Figure 4.1 gives a schematic representation which may help to clarify the assumptions here. The "horizontal" sheet represents the model, while the "vertical" sheet represents the set of all possible true distributions under our assumptions. Clearly, the intersection of these is the set of distributions under the null hypothesis. To further illustrate this we now give two examples.

Example 1 - Normal Distribution

If we have a normal model with density

$$f(x;\mu,\sigma) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2}, \tag{4.4}$$

where $-\infty < x, \mu < \infty$ and $\sigma > 0$, then the Fraser information, defined by equation (4.1), is

$$F(\mu,\sigma) = \int_{-\infty}^{\infty} \left\{ -\frac{1}{2}\log(2\pi) - \log\sigma - \frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2 \right\} g(x)\, dx. \tag{4.5}$$

If we wish to test the null hypothesis $H_0 : \mu = 0$ against the alternative hypothesis $H_1 : \mu \neq 0$, then we assume that the true density, $g(x)$, is such that the maximum of the Fraser information occurs when $\mu = 0$. Thus we consider

$$\frac{\partial F}{\partial \mu}(\mu,\sigma) = \int_{-\infty}^{\infty} \frac{1}{\sigma}\left(\frac{x-\mu}{\sigma}\right) g(x)\, dx, \tag{4.6}$$

$$\frac{\partial F}{\partial \sigma}(\mu,\sigma) = \int_{-\infty}^{\infty} \left\{ -\frac{1}{\sigma} + \frac{(x-\mu)^2}{\sigma^3} \right\} g(x)\, dx. \tag{4.7}$$

We require that

$$\frac{\partial F}{\partial \mu}(0,\sigma) = 0 \tag{4.8}$$

$$\frac{\partial F}{\partial \sigma}(0,\sigma) = 0. \tag{4.9}$$

Substituting (4.6) and (4.7) into (4.8) and (4.9) respectively gives the conditions on $g(x)$ as

$$\int_{-\infty}^{\infty} \frac{x}{\sigma^2} g(x)\, dx = 0 \tag{4.10}$$

$$\int_{-\infty}^{\infty} \left\{ \frac{x^2 - \sigma^2}{\sigma^3} \right\} g(x)\, dx = 0. \tag{4.11}$$

Equation (4.11) is a condition in terms of the unknown variance of the normal model, which is a nuisance parameter here, so it does not in fact place any restriction on the form of $g(x)$. Equation (4.10) requires that the density $g(x)$ has mean 0, which is an intuitively reasonable condition here.

Example 2 - A Four Parameter Directional Distribution

Consider the four-parameter circular distribution, discussed in detail in Yfantis & Borgman (1982), with density

$$f(x;a,b,\alpha,\beta) = [c(a,b,\alpha,\beta)]^{-1} \exp\left\{a\cos(x-\alpha) + b\cos 2(x-\beta)\right\}, \tag{4.12}$$

where $0 \leq \alpha, \beta, x < 2\pi$ and $a, b \geq 0$. Yfantis and Borgman show that the constant $c(a, b, \alpha, \beta)$ is given by

$$c(a,b,\alpha,\beta) = 2\pi \left\{ I_0(a)I_0(b) + 2\sum_{n=1}^{\infty} I_n(b)I_{2n}(a)\cos 2n(\beta - \alpha) \right\}, \tag{4.13}$$

where $I_n(a)$, $I_n(b)$ are modified Bessel functions of the first kind and order n, defined by

$$I_n(s) = \frac{1}{\pi}\int_0^{\pi} e^{s\cos\theta}\cos n\theta \, d\theta. \tag{4.14}$$

Suppose that under this model we wish to test the null hypothesis $H_0 : a = 0$ against the alternative $H_1 : a \neq 0$, i.e. we are testing whether the model can be reduced to the Doubled Von Mises distribution. For this example, the Fraser information is

$$F(a,b,\alpha,\beta) = \int_0^{2\pi} \{-\log c(a,b,\alpha,\beta) + a\cos(x-\alpha) + b\cos 2(x-\beta)\}\, g(x)\, dx. \tag{4.15}$$

To maximize $F(a, b, \alpha, \beta)$ over the parameters consider

$$\begin{aligned}
\frac{\partial F}{\partial a} &= \frac{\partial}{\partial a}\int_0^{2\pi} \{-\log c(a,b,\alpha,\beta) + a\cos(x-\alpha) + b\cos 2(x-\beta)\}\, g(x)\, dx \\
&= \int_0^{2\pi} \left\{ -\frac{\frac{\partial c}{\partial a}(a,b,\alpha,\beta)}{c(a,b,\alpha,\beta)} + \cos(x-\alpha) \right\} g(x)\, dx, && (4.16) \\
\frac{\partial F}{\partial b} &= \int_0^{2\pi} \left\{ -\frac{\frac{\partial c}{\partial b}(a,b,\alpha,\beta)}{c(a,b,\alpha,\beta)} + \cos 2(x-\beta) \right\} g(x)\, dx, && (4.17) \\
\frac{\partial F}{\partial \alpha} &= \int_0^{2\pi} \left\{ -\frac{\frac{\partial c}{\partial \alpha}(a,b,\alpha,\beta)}{c(a,b,\alpha,\beta)} + a\sin(x-\alpha) \right\} g(x)\, dx, && (4.18) \\
\frac{\partial F}{\partial \beta} &= \int_0^{2\pi} \left\{ -\frac{\frac{\partial c}{\partial \beta}(a,b,\alpha,\beta)}{c(a,b,\alpha,\beta)} + 2b\sin 2(x-\beta) \right\} g(x)\, dx. && (4.19) \\
& && (4.20)
\end{aligned}$$

We require that $g(x)$ satisfies

$$\begin{aligned}
\frac{\partial F}{\partial a}(0,b,\alpha,\beta) &= 0, && (4.21) \\
\frac{\partial F}{\partial b}(0,b,\alpha,\beta) &= 0, && (4.22) \\
\frac{\partial F}{\partial \alpha}(0,b,\alpha,\beta) &= 0, && (4.23) \\
\frac{\partial F}{\partial \beta}(0,b,\alpha,\beta) &= 0. && (4.24)
\end{aligned}$$

Noting that

$$c(0, b, \alpha, \beta) = 2\pi I_0(b) \tag{4.25}$$

and that

$$\begin{aligned}
\frac{\partial c}{\partial a}(0, b, \alpha, \beta) &= 0, &(4.26)\\
\frac{\partial c}{\partial b}(0, b, \alpha, \beta) &= 2\pi I_1(b), &(4.27)\\
\frac{\partial c}{\partial \alpha}(0, b, \alpha, \beta) &= 0, &(4.28)\\
\frac{\partial c}{\partial \beta}(0, b, \alpha, \beta) &= 0, &(4.29)
\end{aligned}$$

if we substitute (4.16) - (4.19) into (4.21) - (4.24), we find that the conditions on $g(x)$ are

$$\begin{aligned}
\int_0^{2\pi} \cos(x-\alpha) g(x)\, dx &= 0, &(4.30)\\
\int_0^{2\pi} \left\{ \frac{I_1(b)}{I_0(b)} + \cos 2(x-\beta) \right\} g(x)\, dx &= 0, &(4.31)\\
\int_0^{2\pi} 0 \; g(x)\, dx &= 0, &(4.32)\\
\int_0^{2\pi} 2b \sin 2(x-\beta) g(x)\, dx &= 0. &(4.33)
\end{aligned}$$

Condition (4.32) is clearly trivial, and conditions (4.31) and (4.33) are in terms of nuisance parameters b and β, which are unspecified and therefore do not result in any restriction on $g(x)$. To understand condition (4.30) we write $g(x)$ in terms of its Fourier expansion

$$g(x) = \frac{1}{2\pi} \left[1 + 2 \sum_{p=1}^{\infty} (C_p \cos px + S_p \sin px) \right], \tag{4.34}$$

where $| C_p |, | S_p | \leq 1$. Equation (4.30) then becomes

$$\begin{aligned}
\int_0^{2\pi} \cos(x-\alpha) \frac{1}{2\pi} \left[1 + 2 \sum_{p=1}^{\infty} (C_p \cos px + S_p \sin px) \right] dx &= 0,\\
\frac{1}{2\pi} \sum_{p=1}^{\infty} \left\{ \int_0^{2\pi} C_p \cos\left[(p+1)x - \alpha\right] dx \right.&\\
+ \int_0^{2\pi} C_p \cos\left[(p-1)x + \alpha\right] dx&\\
+ \int_0^{2\pi} S_p \sin\left[(p+1)x - \alpha\right] dx&\\
\left. + \int_0^{2\pi} S_p \sin\left[(p-1)x + \alpha\right] dx \right\} &= 0\\
\frac{1}{2\pi} \left\{ \int_0^{2\pi} C_1 \cos\alpha\, dx + \int_0^{2\pi} S_1 \sin\alpha\, dx \right\} &= 0. &(4.35)
\end{aligned}$$

Again, this is an intuitively reasonable condition on $g(x)$.

4.1 The Likelihood Ratio Test Statistic

4.1.1 Asymptotic Expansion

Firstly, we seek an asymptotic expansion for the likelihood ratio statistic, w, in the mis-specified case, making the assumption given in (4.1) - (4.3). As in the null case, we use Taylor expansion to express w in terms of the joint cumulants of log-likelihood derivatives. As we do not assume that a true value of θ exists, we take Taylor expansions about $\theta(g)$. Partitioning θ into a p-dimensional interest parameter ψ and a q-dimensional nuisance parameter λ, with $p+q=m$, as before, we can see that if we write the null hypothesis as

$$H_0 : \psi = \psi_0, \tag{4.36}$$

where ψ_0 is some specified value, then

$$\theta(g)^T = \left(\psi_0^T, \lambda^T\right). \tag{4.37}$$

For brevity we may write $\left(\psi_0^T, \lambda^T\right) = \theta_0^T$.

Analogously to the null case, let the derivatives of the log likelihood function under H_1, evaluated at θ_0, be written as

$$\begin{aligned} U_r &= \frac{\partial l}{\partial \theta^r}(\theta_0; x), && (4.38) \\ U_{rs} &= \frac{\partial^2 l}{\partial \theta^r \partial \theta^s}(\theta_0; x), && (4.39) \\ U_{rst} &= \frac{\partial^3 l}{\partial \theta^r \theta^s \theta^t}(\theta_0; x), && (4.40) \end{aligned}$$

and so on, and let the joint cumulants be

$$\begin{aligned} \kappa_r &= \mathrm{E}_g[U_r] = 0, && (4.41) \\ \kappa_{r,s} &= \mathrm{E}_g[U_r U_s], && (4.42) \\ \kappa_{rs} &= \mathrm{E}_g[U_{rs}], && (4.43) \\ \kappa_{rs,tu} &= \mathrm{E}_g[U_{rs} U_{tu}] - \kappa_{rs}\kappa_{tu}, && (4.44) \end{aligned}$$

and so on, where E_g denotes expectation under the true density $g(x)$, i.e.

$$\mathrm{E}_g\left[\frac{\partial l}{\partial \theta_r}(\theta_0; x)\right] = \int \frac{\partial l}{\partial \theta_r}(\theta_0; x) g(x)\, dx. \tag{4.45}$$

It is important to note that, for a mis-specified model,

$$\kappa_{rs} \neq -\kappa_{r,s}. \tag{4.46}$$

To see this, consider the derivation of the result that $\kappa_{rs} = -\kappa_{r,s}$ in the null case. We start with the density function, $f(x;\theta_0)$, at the true value θ_0 of θ. By honesty,

$$\int f(x;\theta_0)\,dx = 1, \tag{4.47}$$

and differentiating both sides with respect to θ_0 gives

$$\begin{aligned} \frac{\partial}{\partial\theta_0}\left\{\int f(x;\theta_0)\,dx\right\} &= 0 \\ \int \frac{\partial}{\partial\theta_0}\left\{e^{l(\theta_0;x)}\right\} dx &= 0 \\ \int \frac{\partial l(\theta_0;x)}{\partial\theta_0} f(x;\theta_0)\,dx &= 0. \end{aligned} \tag{4.48}$$

By differentiating again with respect to θ_0 we have

$$\begin{aligned} \frac{\partial}{\partial\theta_0^T}\left\{\int \frac{\partial l(\theta_0;x)}{\partial\theta_0} f(x;\theta_0)\,dx\right\} &= 0 \\ \int \left\{\frac{\partial^2 l(\theta_0;x)}{\partial\theta_0\partial\theta_0^T} + \frac{\partial l(\theta_0;x)}{\partial\theta_0}\frac{\partial l(\theta_0;x)}{\partial\theta_0^T}\right\} f(x;\theta_0)\,dx &= 0. \end{aligned} \tag{4.49}$$

Specifically, if we consider the (r,s)-entries in the matrices in equation (4.49), we have

$$\begin{aligned} \int \frac{\partial^2 l(\theta_0;x)}{\partial\theta_0^r\partial\theta_0^s} f(x;\theta_0\,dx &= -\int \frac{\partial l(\theta_0;x)}{\partial\theta_0^r}\frac{\partial l(\theta_0;x)}{\partial\theta_0^s} f(x;\theta_0)\,dx \\ \kappa_{rs} &= -\kappa_{r,s}. \end{aligned} \tag{4.50}$$

Further differentiation leads to similar identities for higher-order cumulants.

Under a mis-specified model, if $\theta_0 = \theta(g)$, then from equations (4.1) and (4.2) we have

$$\int \frac{\partial l(\theta_0;x)}{\partial\theta_0} g(x)\,dx = 0. \tag{4.51}$$

There is, however, no way of differentiating (4.51) to obtain identities similar to those under the null model, so we do not have $\kappa_{rs} = -\kappa_{r,s}$, nor in fact do any of the null identities hold. This creates the fundamental difference between the asymptotic distributions of the statistics we consider under the null and mis-specified models.

Analogously to the null case, we define the matrices H and K by

$$\begin{aligned} H &= (\kappa_{ij})_{m\times m} \\ &= \begin{bmatrix} H_{\psi\psi} & H_{\psi\lambda} \\ H_{\lambda\psi} & H_{\lambda\lambda} \end{bmatrix}, \qquad (4.52) \\ K &= (\kappa_{i,j})_{m\times m} \\ &= \begin{bmatrix} K_{\psi,\psi} & K_{\psi,\lambda} \\ K_{\lambda,\psi} & K_{\lambda,\lambda} \end{bmatrix}, \qquad (4.53) \end{aligned}$$

and note that we do not have $-H = K = I$, where I is the Fisher information matrix, as we did under the null model. Taylor expansion of the likelihood equations

$$\frac{\partial l}{\partial \theta^r}(\hat{\theta}; x) = 0 \qquad (4.54)$$

about $\theta = \theta_0$ gives

$$0 = U_r + U_{rs}(\hat{\theta} - \theta_0)^s + \frac{1}{2!}U_{rst}(\hat{\theta} - \theta_0)^s(\hat{\theta} - \theta_0)^t + \ldots. \qquad (4.55)$$

We can make the dependence on n explicit by writing

$$U_r = n^{\frac{1}{2}} Z_r, \qquad (4.56)$$

$$U_{rs} = n\kappa_{rs} + n^{\frac{1}{2}} Z_{rs}, \qquad (4.57)$$

$$U_{rst} = n\kappa_{rst} + n^{\frac{1}{2}} Z_{rst}, \qquad (4.58)$$

and so on, and

$$\hat{\delta} = n^{\frac{1}{2}}(\hat{\theta} - \theta_0), \qquad (4.59)$$

so that we get

$$\begin{aligned} 0 &= n^{\frac{1}{2}}(Z_r + \kappa_{rs}\hat{\delta}^s) \\ &\quad + \left(Z_{rs}\hat{\delta}^s + \frac{1}{2}\kappa_{rst}\hat{\delta}^s\hat{\delta}^t\right) \\ &\quad + n^{-\frac{1}{2}}\left(\frac{1}{2}Z_{rst}\hat{\delta}^s\hat{\delta}^t + \frac{1}{6}\kappa_{rstu}\hat{\delta}^s\hat{\delta}^t\hat{\delta}^u\right) + O_p(n^{-1}). \qquad (4.60) \end{aligned}$$

Setting the $O_p\left(n^{\frac{1}{2}}\right)$, $O_p(1)$ and $O_p\left(n^{-\frac{1}{2}}\right)$ terms in turn equal to zero and solving for $\hat{\delta}$ gives

$$\hat{\delta}^r = -\kappa^{rs}Z_s + \left(\kappa^{rs}\kappa^{tu}Z_{st}Z_u - \frac{1}{2}\kappa^{rst}Z_sZ_t\right)n^{-\frac{1}{2}} + \ldots, \qquad (4.61)$$

where κ^{rs} is the matrix inverse of κ_{rs}, so that

$$\begin{aligned} \kappa^{rs}\kappa_{st} &= \delta^r_t, \\ \kappa^{rs} &= \kappa^{ri}\kappa^{sj}\kappa_{ij}. \end{aligned}$$

If we write (4.61) as

$$\hat{\delta}^r = -Z^r + c^r n^{-\frac{1}{2}} + d^r n^{-1} + O_p(n^{-\frac{3}{2}}), \tag{4.62}$$

where $Z^r = \kappa^{rs} Z_s$, then Taylor expansion of $l(\hat{\theta}; x)$ gives

$$\begin{aligned} l(\hat{\theta}; x) &= l(\theta_0; x) - \frac{1}{2}\kappa^{rs} Z_r Z_s \\ &\quad + n^{-\frac{1}{2}} \left(\frac{1}{2} Z_{rs} Z^r Z^s - \frac{1}{6}\kappa_{rst} Z^r Z^s Z^t \right) \\ &\quad + n^{-1} \left(\frac{1}{2}\kappa_{rs} c^r c^s - Z_{rs} Z^r c^s + \frac{1}{2}\kappa_{rst} Z^r Z^s c^t \right. \\ &\qquad \left. - \frac{1}{6} Z_{rst} Z^r Z^s Z^t + \frac{1}{24}\kappa_{rstu} Z^r Z^s Z^t Z^u \right) \\ &\quad + O_p(n^{-\frac{3}{2}}). \end{aligned} \tag{4.63}$$

On substituting for c^r we obtain

$$\begin{aligned} l(\hat{\theta}; x) - l(\theta_0; x) &= -\frac{1}{2}\kappa^{rs} Z_r Z_s \\ &\quad + n^{-\frac{1}{2}} \left(\frac{1}{2} Z_{rs} Z^r Z^s - \frac{1}{6}\kappa_{rst} Z^r Z^s Z^t \right) \\ &\quad + n^{-1} \left[-\frac{1}{2} \left(Z_{ri} Z^i - \frac{1}{2}\kappa_{rij} Z^i Z^j \right) \kappa^{rs} \left(Z_{sk} Z^k - \frac{1}{2}\kappa_{skl} Z^k Z^l \right) \right. \\ &\qquad \left. - \frac{1}{6} Z_{rst} Z^r Z^s Z^t + \frac{1}{24}\kappa_{rstu} Z^r Z^s Z^t Z^u \right] \\ &\quad + O_p(n^{-\frac{3}{2}}). \end{aligned} \tag{4.64}$$

By Taylor expansion of the likelihood equations under H_0 we can obtain the analogous expression for $l(\tilde{\theta}; x)$, and hence we have an expression for w. Working with this full expression, however, is cumbersome, so we obtain an expansion now assuming $\dim \lambda = 0$ and show in section 4.1.3 how to obtain the distributional results for the general case from the simplified one.

4.1.2 Distribution of w when $\dim \lambda = 0$

In order to obtain the moment generating function of w, we will write w as a quadratic in a vector $\mathbf{W}$, for which we can fairly easily find the cumulant generating function. In this section, we derive the cumulant generating function of $\mathbf{W}$ under the simplifying assumption that $\dim \lambda = 0$, that is, there is no nuisance parameter present. In the next section it will be shown that the cumulant generating function for $\mathbf{W}$ can easily be generalized to the case where there is a nuisance parameter.

When $\dim \lambda = 0$, that is, there is no nuisance parameter, the likelihood ratio test statistic is $w = 2[l(\hat{\theta}; x) - l(\theta_0; x)]$, so $w/2$ is given by (4.64), i.e.

$$\begin{aligned} w &= -\kappa^{rs} Z_r Z_s \\ &\quad + n^{-\frac{1}{2}} \left(Z_{rs} Z^r Z^s - \frac{1}{3} \kappa_{rst} Z^r Z^s Z^t \right) \\ &\quad + n^{-1} \left[- \left(Z_{ri} Z^i - \frac{1}{2} \kappa_{rij} Z^i Z^j \right) \kappa^{rs} \left(Z_{sk} Z^k - \frac{1}{2} \kappa_{skl} Z^k Z^l \right) \right. \\ &\qquad \left. - \frac{1}{3} Z_{rst} Z^r Z^s Z^t + \frac{1}{12} \kappa_{rstu} Z^r Z^s Z^t Z^u \right] \\ &\quad + O_p\left(n^{-\frac{3}{2}}\right). \end{aligned} \tag{4.65}$$

To ease computation we use, as in the case of the null distribution, a parameterization based on the expected likelihood yoke (see sections 2.1.2 and 2.2.5). Take any fixed value of θ – here we will take θ_0 – then define β^r_{st} and β^r_{stu} by (2.16) and (2.17). We define a parameter transformation in a neighbourhood of θ_0 by (2.18). As in the null case, we will denote the joint cumulants of the log-likelihood derivatives in the ϕ-parameterization by ν_r, ν_{rs}, $\nu_{r,s}$ and so on, and note also that

$$\nu_{r,st} = \nu_{r,stu} = \ldots = 0. \tag{4.66}$$

Working in this parameterization, we can rewrite (4.65) as

$$w = -W_r W_s \nu^{rs}, \tag{4.67}$$

where

$$\begin{aligned} W_r &= Z_r \\ &\quad + n^{-\frac{1}{2}} \left(\frac{1}{6} \nu_{rij} Z^i Z^j - \frac{1}{2} Z_{ri} Z^i \right) \\ &\quad + n^{-1} \left(\frac{1}{6} Z_{rij} Z^i Z^j - \frac{1}{24} \nu_{rijk} Z^i Z^j Z^k + \frac{3}{8} Z_{ri} Z^{ij} Z_j \right. \\ &\qquad \left. - \frac{5}{12} Z_{ri} Z_j Z_k \nu^{ijk} + \frac{1}{9} \nu_{rij} \nu_{klm} \nu^{jk} Z^i Z^l Z^m \right) \\ &\quad + O(n^{-\frac{3}{2}}). \end{aligned} \tag{4.68}$$

Noting that

$$\mathrm{E}(Z_r) = \mathrm{E}(Z_{rs}) = \mathrm{E}(Z_{rst}) = 0,$$

we can find expressions for the joint moments of the Zs, i.e.

$$\mathrm{E}(Z_r Z_s) \;=\; \nu_{r,s},$$

$$
\begin{aligned}
\mathrm{E}(Z_r Z_{st}) &= \nu_{r,st}, \\
\mathrm{E}(Z_r Z_s Z_t) &= n^{-\frac{1}{2}} \nu_{r,s,t}, \\
\mathrm{E}(Z_r Z_s Z_{tu}) &= n^{-\frac{1}{2}} \nu_{r,s,tu}, \\
\mathrm{E}(Z_r Z_s Z_t Z_u) &= n^{-1} \nu_{r,s,t,u} + \nu_{r,s}\nu_{t,u}[3], \\
\mathrm{E}(Z_r Z_s Z_{tu} Z_{vw}) &= n^{-1} \nu_{r,s,tu,vw} + \nu_{r,s}\nu_{tu,vw},
\end{aligned}
$$

etc. Thus we can also show that

$$
\begin{aligned}
\mathrm{E}(W_r) &= n^{-\frac{1}{2}} \frac{1}{6} \nu_{rij}\nu^{ia}\nu^{jb}\nu_{a,b} + O(n^{-\frac{3}{2}}), && (4.69) \\
\mathrm{Cov}(W_r, W_s) &= \nu_{r,s} \\
&\quad + n^{-1}\Big(\frac{1}{3}\nu_{rij}\nu^{ia}\nu^{jb}\nu_{s,a,b} \\
&\qquad - \nu^{ia}\nu_{s,a,ri} - \frac{1}{4}\nu_{rijk}\nu^{ia}\nu^{jb}\nu^{kc}\nu_{s,a}\nu_{b,c} \\
&\qquad + \frac{3}{4}\nu^{ia}\nu^{jb}\nu_{s,j}\nu_{ri,ab} + \frac{1}{4}\nu^{ia}\nu^{\alpha b}\nu_{ri,s\alpha}\nu_{a,b} \\
&\qquad + \frac{2}{9}\nu_{rij}\nu_{klm}\nu^{jk}\nu^{ia}\nu^{lb}\nu^{mc}\nu_{s,a}\nu_{b,c} \\
&\qquad + \frac{4}{9}\nu_{rij}\nu_{klm}\nu^{jk}\nu^{ia}\nu^{lb}\nu^{mc}\nu_{s,b}\nu_{a,c} \\
&\qquad + \frac{1}{18}\nu_{rij}\nu^{ia}\nu^{jb}\nu_{s\alpha\beta}\nu^{\alpha c}\nu^{\beta d}\nu_{a,c}\nu_{b,d}\Big) \\
&\quad + O(n^{-\frac{3}{2}}), && (4.70) \\
\mathrm{Cum}(W_r, W_s, W_t) &= n^{-\frac{1}{2}}\left[\nu_{r,s,t} + \nu_{tij}\nu^{ia}\nu^{jb}\nu_{r,a}\nu_{s,b}\right] + O(n^{-\frac{3}{2}}) && (4.71) \\
\mathrm{Cum}(W_r, W_s, W_t, W_u) &= n^{-1}\Big[\nu_{r,s,t,u} + 4\nu_{uij}\nu^{ia}\nu^{jb}(\nu_{a,r}\nu_{b,s,t}) \\
&\qquad - 6\nu^{ia}\nu_{a,r}\nu_{s,t,ui} - \nu_{uijk}\nu^{ia}\nu^{jb}\nu^{kc}\nu_{r,a}\nu_{s,b}\nu_{t,c} \\
&\qquad + \frac{8}{3}\nu_{uij}\nu_{klm}\nu^{jk}\nu^{ia}\nu^{lb}\nu^{mc}\nu_{r,a}\nu_{s,b}\nu_{t,c} \\
&\qquad + \frac{4}{3}\nu_{uij}\nu^{ia}\nu^{jb}\nu_{tkl}\nu^{kc}\nu^{ld}\nu_{r,a}\nu_{s,c}\nu_{b,d} \\
&\qquad + 3\nu^{ia}\nu^{jb}\nu_{ti,uj}\nu_{r,a}\nu_{s,b}\Big] \\
&\quad + O(n^{-\frac{3}{2}}) && (4.72)
\end{aligned}
$$

and higher order joint cumulants are of order $O(n^{-\frac{3}{2}})$ or smaller. Thus, up to terms of order $O(n^{-1})$, we can write the cumulants of $\boldsymbol{W} = (W_1, \ldots, W_p)^T$ as

$$
\begin{aligned}
\mathrm{E}(W_r) &= n^{-\frac{1}{2}}\delta_r, \\
\mathrm{Cov}(W_r, W_s) &= C_{r,s} + n^{-1}D_{r,s} \\
\mathrm{Cum}(W_r, W_s, W_t) &= n^{-\frac{1}{2}}F_{r,s,t}
\end{aligned}
$$

$$\mathrm{Cum}(W_r, W_s, W_t, W_u) \;=\; n^{-1}G_{r,s,t,u},$$

where δ_r, $C_{r,s}$, $D_{r,s}$, $F_{r,s,t}$ and $G_{r,s,t,u}$ are defined by (4.69) - (4.72) above. Note that we may obtain the expressions for the null cumulants from (4.69) - (4.72) by making the substitution $g(x) = f(x;\theta_0)$, and noting that in this case, $\nu_{rs} = -\nu_{r,s}$. In the null case

$$F_{r,s,t} = G_{r,s,t,u} = 0, \tag{4.73}$$

and hence

$$\mathbf{W} \sim MVN(n^{-\frac{1}{2}}\boldsymbol{\delta}, C + n^{-1}D), \tag{4.74}$$

neglecting terms of order $O(n^{-\frac{3}{2}})$ or smaller, where the components of the mean vector $n^{-\frac{1}{2}}\boldsymbol{\delta}$ and the covariance matrix $C + n^{-1}D$ are given above. This is the distributional result which McCullagh (1987) uses in the null case to show that Bartlett adjustment works (see section 2.1.2).

We can return to our original parameterization by expressing the νs in terms of the κs. We find that, for the mis-specified model,

$$\begin{aligned}
\nu_{r,s} &= \kappa_{r,s} && (4.75)\\
\nu_{rs} &= \kappa_{rs} && (4.76)\\
\nu_{r,s,t} &= \kappa_{r,s,t} && (4.77)\\
\nu_{rst} &= \kappa_{rst} - \kappa^{\alpha,\beta}\kappa_{\beta,rs}\kappa_{\alpha t}[3] && (4.78)\\
\nu_{r,s,t,u} &= \kappa_{r,s,t,u} && (4.79)\\
\nu_{r,s,tu} &= \kappa_{r,s,tu} - \kappa^{\alpha,\beta}\kappa_{\beta,tu}\kappa_{r,s,\alpha} && (4.80)\\
\nu_{rs,tu} &= \kappa_{rs,tu} - \kappa^{\alpha,\beta}\kappa_{\beta,rs}\kappa_{\alpha,tu} && (4.81)\\
\nu_{rstu} &= \kappa_{rstu} - \kappa^{\alpha,\beta}\kappa_{\beta,rs}\kappa_{\alpha tu}[6] - \kappa^{\alpha,\beta}\kappa_{\beta,rs}\left(\kappa^{\gamma,\delta}\kappa_{\delta,\alpha t}\kappa_{\gamma u}[3]\right)[6], && (4.82)
\end{aligned}$$

and substituting into (4.69) - (4.72) gives

$$\begin{aligned}
\delta_r &= \frac{1}{6}\left[\left(\kappa_{rij} - \kappa^{\alpha,\beta}\kappa_{\beta,ri}\kappa_{\alpha j}[3]\right)\kappa^{ia}\kappa^{jb}\kappa_{a,b}\right] && (4.83)\\
C_{r,s} &= \kappa_{r,s}\\
D_{r,s} &= \frac{1}{3}\left(\kappa_{rij} - \kappa^{\alpha,\beta}\kappa_{\beta,ri}\kappa_{\alpha j}[3]\right)\kappa^{ia}\kappa^{jb}\kappa_{s,a,b} - \kappa^{ia}\left(\kappa_{s,a,ri} - \kappa^{\alpha,\beta}\kappa_{\beta,ri}\kappa_{s,a,\alpha}\right)\\
&\quad -\frac{1}{4}\left\{\kappa_{rijk} - \kappa^{\alpha,\beta}\kappa_{\beta,ri}\kappa_{\alpha jk}[6] - \kappa^{\alpha,\beta}\kappa_{\beta,ri}\left(\kappa^{\gamma,\delta}\kappa_{\delta,\alpha j}\kappa_{\gamma k}[3]\right)[6]\right\}\\
&\qquad \kappa^{ia}\kappa^{jb}\kappa^{kc}\kappa_{s,a}\kappa_{b,c}
\end{aligned}$$

$$
\begin{aligned}
&+\frac{3}{4}\kappa^{ia}\kappa^{jb}\kappa_{s,j}\left(\kappa_{ri,ab}-\kappa^{\alpha,\beta}\kappa_{\beta,ri}\kappa_{\alpha,ab}\right)\\
&+\frac{1}{4}\kappa^{ia}\kappa^{\alpha b}\kappa_{a,b}\left(\kappa_{ri,s\alpha}-\kappa^{\beta,\gamma}\kappa_{\gamma,ri}\kappa_{\beta,s\alpha}\right)\\
&+\left(\kappa_{rij}-\kappa^{\alpha,\beta}\kappa_{\beta,ri}\kappa_{\alpha j}[3]\right)\left(\kappa_{klm}-\kappa^{\gamma,\delta}\kappa_{\delta,kl}\kappa_{\gamma m}[3]\right)\kappa^{jk}\kappa^{ia}\kappa^{lb}\kappa^{mc}\\
&\qquad\left(\frac{2}{9}\kappa_{s,a}\kappa_{b,c}+\frac{4}{9}\kappa_{s,b}\kappa_{a,c}\right)\\
&+\frac{1}{18}\left(\kappa_{rij}-\kappa^{\gamma,\delta}\kappa_{\delta,ri}\kappa_{\gamma j}[3]\right)\left(\kappa_{s\alpha\beta}-\kappa^{\epsilon,\theta}\kappa_{\theta,s\alpha}\kappa_{\epsilon\beta}[3]\right)\\
&\qquad\kappa^{ia}\kappa^{jb}\kappa^{\alpha c}\kappa^{\beta d}\kappa_{a,c}\kappa_{b,d}. \qquad (4.84)
\end{aligned}
$$

$$
F_{r,s,t} = \kappa_{r,s,t}+\left(\kappa_{tij}-\kappa^{\alpha,\beta}\kappa_{\beta,ti}\kappa_{\alpha j}[3]\right)\kappa^{ia}\kappa^{jb}\kappa_{r,a}\kappa_{s,b}, \qquad (4.85)
$$

$$
\begin{aligned}
G_{r,s,t,u} = {} & \kappa_{r,s,t,u}+4\left(\kappa_{uij}-\kappa^{\alpha,\beta}\kappa_{\beta,ui}\kappa_{\alpha j}[3]\right)\kappa^{ia}\kappa^{jb}\kappa_{a,r}\kappa_{b,s,t}\\
&-6\kappa^{ia}\kappa_{a,r}\left(\kappa_{s,t,ui}-\kappa^{\alpha,\beta}\kappa_{\beta,ui}\kappa_{\alpha,s,t}\right)\\
&-\Big\{\kappa_{uijk}-\kappa^{\alpha,\beta}\kappa_{\beta,ui}\kappa_{\alpha jk}[6]-\kappa^{\alpha,\beta}\kappa_{\beta,ui}\left(\kappa^{\gamma,\delta}\kappa_{\delta,\alpha j}\kappa_{\gamma k}[3]\right)[6]\\
&\qquad-\kappa^{\alpha,\beta}\kappa_{\beta,ui}\kappa^{\gamma,\delta}\kappa_{\delta,jk}\kappa_{\alpha\gamma}[3]-\kappa^{\alpha,\beta}\kappa_{\beta,uij}\kappa_{\alpha k}[4]\Big\}\\
&\qquad\qquad\kappa^{ia}\kappa^{jb}\kappa^{kc}\kappa_{r,a}\kappa_{s,b}\kappa_{t,c}\\
&+\frac{8}{3}\left(\kappa_{uij}-\kappa^{\alpha,\beta}\kappa_{\beta,ui}\kappa_{\alpha j}[3]\right)\left(\kappa_{klm}-\kappa^{\gamma,\delta}\kappa_{\delta,kl}\kappa_{\gamma m}[3]\right)\\
&\qquad\kappa^{jk}\kappa^{ia}\kappa^{lb}\kappa^{mc}\kappa_{r,a}\kappa_{s,b}\kappa_{t,c}\\
&+\frac{4}{3}\left(\kappa_{uij}-\kappa^{\alpha,\beta}\kappa_{\beta,ui}\kappa_{\alpha j}[3]\right)\left(\kappa_{tkl}-\kappa^{\gamma,\delta}\kappa_{\delta,tk}\kappa_{\gamma l}[3]\right)\\
&\qquad\kappa^{ia}\kappa^{jb}\kappa^{kc}\kappa^{ld}\kappa_{r,a}\kappa_{s,c}\kappa_{b,d}\\
&+3\kappa^{ia}\kappa^{jb}\left(\kappa_{ti,uj}-\kappa^{\alpha,\beta}\kappa_{\beta,ti}\kappa_{\alpha,uj}\right)\kappa_{r,a}\kappa_{s,b}. \qquad (4.86)
\end{aligned}
$$

We can write equation (4.67) in matrix notation as

$$
w=-\boldsymbol{W}^TH^{-1}\boldsymbol{W},
$$

where matrix H is given by equation (4.52). The cumulant generating function of $\boldsymbol{W}$, $K_{\mathbf{W}}(t)$, is given by

$$
\begin{aligned}
K_{\mathbf{W}}(t) &= n^{-\frac{1}{2}}\delta_r t^r+\frac{1}{2!}(C_{r,s}+n^{-1}D_{r,s})t^r t^s\\
&\quad+\frac{1}{3!}n^{-\frac{1}{2}}F_{r,s,t}t^r t^s t^t+\frac{1}{4!}n^{-1}G_{r,s,t,u}t^r t^s t^t t^u+O(n^{-\frac{3}{2}})\\
&= \frac{1}{2}C_{r,s}t^r t^s+n^{-\frac{1}{2}}\left(\delta_r t^r+\frac{1}{6}F_{r,s,t}t^r t^s t^t\right)\\
&\quad+n^{-1}\left(\frac{1}{2}D_{r,s}t^r t^s+\frac{1}{24}G_{r,s,t,u}t^r t^s t^t t^u\right)+O(n^{-\frac{3}{2}}). \qquad (4.87)
\end{aligned}
$$

4.1.3 Distribution of w when $\dim \lambda \neq 0$

Now consider the general case, where there is a nuisance parameter present. We will show here that the cumulant generating function for $\mathbf{W}$ in this case may be derived quite simply from that given in equation (4.87), where there is no nuisance parameter.

We can see that the likelihood ratio statistic (2.1) can be written as

$$w = -\boldsymbol{V}^T M^{-1} \boldsymbol{V} \tag{4.88}$$

where

$$\boldsymbol{V} = \boldsymbol{W}_\psi - H_{\psi\lambda} H_{\lambda\lambda}^{-1} \boldsymbol{W}_\lambda, \tag{4.89}$$

$\mathbf{W}$ is partitioned in the obvious way as

$$\boldsymbol{W} = [\boldsymbol{W}_\psi^T, \boldsymbol{W}_\lambda^T]^T \tag{4.90}$$

and

$$\begin{aligned} M &= H_{\psi\psi.\lambda} \\ &= H_{\psi\psi} - H_{\psi\lambda} H_{\lambda\lambda}^{-1} H_{\lambda\psi}, \end{aligned} \tag{4.91}$$

with H as defined by (4.52). As in the null case, we definine the matrix B as

$$B = \begin{bmatrix} 0 & 0 \\ 0 & H_{\lambda\lambda}^{-1} \end{bmatrix}, \tag{4.92}$$

and denote the (r, s)-entry of B by b^{rs}. From (4.89) it follows that, up to terms of order $O(n^{-1})$, the cumulants of $\mathbf{V}$ (using index notation for partitioned parameters) are given by

$$\begin{aligned} \mathrm{E}(V_{\psi_1}) &= \delta_{\psi_1} - \kappa_{\psi_1\lambda_1} b^{\lambda_1\lambda_2} \delta_{\lambda_2}, & (4.93) \\ \mathrm{Cov}(V_{\psi_1}, V_{\psi_2}) &= C_{\psi_1,\psi_2} + n^{-1} D_{\psi_1,\psi_2} \\ &\quad -2(C_{\psi_1,\lambda_1} + n^{-1} D_{\psi_1,\lambda_1}) b^{\lambda_1\lambda_2} \kappa_{\lambda_2\psi_2} \\ &\quad +(C_{\lambda_1,\lambda_2} + n^{-1} D_{\lambda_1,\lambda_2}) b^{\lambda_1\lambda_3} b^{\lambda_2\lambda_4} \kappa_{\lambda_3\psi_1} \kappa_{\lambda_4\psi_2} & (4.94) \\ \mathrm{Cum}(V_{\psi_1}, V_{\psi_2}, V_{\psi_3}) &= F_{\psi_1,\psi_2,\psi_3} \\ &\quad -3F_{\psi_1,\psi_2,\lambda_1} b^{\lambda_1\lambda_2} \kappa_{\lambda_2\psi_3} \\ &\quad +3F_{\psi_1,\lambda_1,\lambda_2} b^{\lambda_1\lambda_3} b^{\lambda_2\lambda_4} \kappa_{\lambda_3\psi_2} \kappa_{\lambda_4\psi_3} \end{aligned}$$

$$
-F_{\lambda_1,\lambda_2,\lambda_3} b^{\lambda_1\lambda_4} b^{\lambda_2\lambda_5} b^{\lambda_3\lambda_6} \kappa_{\lambda_4\psi_1} \kappa_{\lambda_5\psi_2} \kappa_{\lambda_6\psi_3}, \quad (4.95)
$$

$$
\begin{aligned}
\mathrm{Cum}(V_{\psi_1}, V_{\psi_2}, V_{\psi_3}, V_{\psi_4}) = \; & G_{\psi_1,\psi_2,\psi_3,\psi_4} \\
& -4G_{\psi_1,\psi_2,\psi_3,\lambda_1} b^{\lambda_1\lambda_2} \kappa_{\lambda_2\psi_4} \\
& +6G_{\psi_1,\psi_2,\lambda_1,\lambda_2} b^{\lambda_1\lambda_3} b^{\lambda_2\lambda_4} \kappa_{\lambda_3\psi_3} \kappa_{\lambda_4\psi_4} \\
& -4G_{\psi_1,\lambda_1,\lambda_2,\lambda_3} b^{\lambda_1\lambda_4} b^{\lambda_2\lambda_5} b^{\lambda_3\lambda_6} \kappa_{\lambda_4\psi_2} \kappa_{\lambda_5\psi_3} \kappa_{\lambda_6\psi_4} \\
& +G_{\lambda_1,\lambda_2,\lambda_3,\lambda_4} b^{\lambda_1\lambda_5} b^{\lambda_2\lambda_6} b^{\lambda_3\lambda_7} b^{\lambda_4\lambda_8} \\
& \qquad \kappa_{\lambda_5\psi_1} \kappa_{\lambda_6\psi_2} \kappa_{\lambda_7\psi_3} \kappa_{\lambda_8\psi_4}. \quad (4.96)
\end{aligned}
$$

Thus we can obtain the most general form of the cumulants, and hence of the cumulant generating function, of $\mathbf{V}$ from those of $\mathbf{W}$ by using (4.93) - (4.96). For simplicity we continue this chapter assuming that dim $\lambda = 0$ but note that the corresponding equations for dim $\lambda \neq 0$ may be obtained by replacing δ_r, $C_{r,s}$, $D_{r,s}$, $F_{r,s,t}$ and $G_{r,s,t,u}$, which define the cumulants of $\mathbf{W}$, by the equivalent expressions from (4.93) - (4.96). The results regarding Bartlett adjustment of the likelihood ratio statistic under a mis-specified model therefore hold for a model with nuisance parameter of arbitrary dimension.

4.1.4 Bartlett Adjustment when dim $\psi = 1$

Kent (1982) shows by means of considering the first order terms of a Taylor expansion that, in the case of a mis-specified model, the likelihood ratio statistic is asymptotically distributed as a weighted sum of independent chi-squared variates. Specifically, if $U_1, \ldots, U_p$ are independent χ^2_1 variates, then, asymptotically as $n \to \infty$

$$
w \sim \sum_{i=1}^{p} \mu_i U_i, \quad (4.97)
$$

where the μ_i are the eigenvalues of the matrix

$$
H_{\psi\psi.\lambda}(H^{-1}KH^{-1})_{\psi\psi}. \quad (4.98)
$$

Note that when the model is not mis-specified we have $-H = K = I$, the Fisher information matrix, and in this case the matrix (4.98) reduces to the identity.

Motivated by this we can now seek to perform Bartlett adjustment, but here require that the distribution of the corrected likelihood ratio test statistic be a weighted sum of chi-squared variates instead of a chi-squared distribution. Thus we aim to find b such that

$$\left(1+\frac{b}{n}\right)^{-1} w \sim \sum_{i=1}^{p} \mu_i U_i$$

with error $O(n^{-\frac{3}{2}})$ or smaller.

Considering the moment generating function of the corrected test statistic $w' = \left(1+\frac{b}{n}\right)^{-1} w$, it can be seen that we require

$$M_{(1+b/n)^{-1}w}(s) = \prod_{i=1}^{p} (1-2\mu_i s)^{-\frac{1}{2}} + O(n^{-\frac{3}{2}}). \tag{4.99}$$

We need to compare the expression (4.99) with the actual moment generating function obtained from the Taylor expansion of w to find out whether Bartlett adjustment is possible, and if so to obtain an expression for b.

Obtaining the moment generating function of w from the cumulant generating function of $\mathbf{W}$, given by (4.87), is algebraically complicated. To ease calculation, we consider just the case of a one-dimensional interest parameter, i.e. the case where $\mathbf{W}$ is in fact a scalar. We may then write (4.87) as

$$K_W(t) = \frac{1}{2}Ct^2 + n^{-\frac{1}{2}}\left(\delta t + \frac{1}{6}Ft^3\right) + n^{-1}\left(\frac{1}{2}Dt^2 + \frac{1}{24}Gt^4\right) + O\left(n^{-\frac{3}{2}}\right), \tag{4.100}$$

so the moment generating function of the scalar W is given by

$$\begin{aligned}
M_W(t) &= \exp\left\{\frac{1}{2}Ct^2 + n^{-\frac{1}{2}}\left(\delta t + \frac{1}{6}Ft^3\right) + n^{-1}\left(\frac{1}{2}Dt^2 + \frac{1}{24}Gt^4\right)\right\} \\
&\quad + O\left(n^{-\frac{3}{2}}\right) \\
&= \exp\left\{\frac{1}{2}Ct^2\right\}\left\{1 + n^{-\frac{1}{2}}\left(\delta + \frac{1}{6}Ft^2\right)t \right. \\
&\quad \left. + n^{-1}\left[\frac{1}{2}(D+\delta^2)t^2 + (\frac{1}{24}G + \frac{1}{6}\delta F)t^4 + \frac{1}{72}F^2t^6\right]\right\} + O\left(n^{-\frac{3}{2}}\right) \\
&= \sum_{r=0}^{\infty} \frac{1}{r!}\left(\frac{1}{2}\right)^r C^r t^{2r} \\
&\quad + n^{-\frac{1}{2}}\left\{\sum_{r=0}^{\infty} \frac{1}{r!}\left(\frac{1}{2}\right)^r \delta C^r t^{2r+1} + \sum_{r=0}^{\infty} \frac{1}{r!}\left(\frac{1}{2}\right)^r \frac{1}{6} F C^r t^{2r+3}\right\} \\
&\quad + n^{-1}\left\{\sum_{r=0}^{\infty} \frac{1}{r!}\left(\frac{1}{2}\right)^{r+1} (D+\delta^2) C^{2r} t^{2(r+1)} \right. \\
&\quad\quad + \sum_{r=0}^{\infty} \frac{1}{r!}\left(\frac{1}{2}\right)^r \left(\frac{1}{24}G + \frac{1}{6}\delta F\right) C^r t^{2(r+2)} \\
&\quad\quad \left. + \sum_{r=0}^{\infty} \frac{1}{r!}\left(\frac{1}{2}\right)\frac{1}{72} F^2 C^r t^{2(r+3)}\right\} \\
&\quad + O\left(n^{-\frac{3}{2}}\right).
\end{aligned} \tag{4.101}$$

From this we may obtain the moment generating function of $w = -WH^{-1}W$,

$$
\begin{aligned}
M_w(s) &= \mathrm{E}[\exp\{ws\}] \\
&= \sum_{r=0}^{\infty} \frac{1}{r!} s^r (-H^{-1})^r \mathrm{E}[W^{2r}] \\
&= \sum_{r=0}^{\infty} \frac{1}{r!} \frac{(2r)!}{r!} \left(\frac{1}{2}\right)^r (-H^{-1}C)^r s^r \\
&\quad + n^{-1} \left\{ \sum_{r=1}^{\infty} \frac{1}{r!} \frac{(2r)!}{(r-1)!} \left(\frac{1}{2}\right)^r (-H^{-1}C)^r C^{-1}(D+\delta^2) s^r \right. \\
&\qquad + \sum_{r=2}^{\infty} \frac{1}{r!} \frac{(2r)!}{(r-2)!} \left(\frac{1}{2}\right)^{r-2} (-H^{-1}C)^r C^{-2} \left(\frac{1}{24}G + \frac{1}{6}\delta F\right) s^r \\
&\qquad \left. + \sum_{r=3}^{\infty} \frac{1}{r!} \frac{(2r)!}{(r-3)!} \left(\frac{1}{2}\right)^{r-3} (-H^{-1}C)^r C^{-3} \frac{1}{72} F^2 s^r \right\} \\
&\quad + O\left(n^{-\frac{3}{2}}\right), \qquad (4.102)
\end{aligned}
$$

and we compare this with the form when $p = 1$ of the moment generating function given by (4.99),

$$
\begin{aligned}
M_{(1+b/n)^{-1}w}(s) &= (1-2\mu s)^{-\frac{1}{2}} + O(n^{-\frac{3}{2}}) \\
M_w(s) &= \left[1 - 2\mu\left(1 + \frac{b}{n}\right)s\right]^{-\frac{1}{2}} + O\left(n^{-\frac{3}{2}}\right) \\
&= (1-2\mu s)^{-\frac{1}{2}} \left[1 - \frac{2\mu s b}{(1-2\mu s)\,n}\right]^{-\frac{1}{2}} + O\left(n^{-\frac{3}{2}}\right) \\
&= (1-2\mu s)^{-\frac{1}{2}} \left[1 + \frac{\mu b s}{n(1-2\mu s)}\right] + O\left(n^{-\frac{3}{2}}\right). \qquad (4.103)
\end{aligned}
$$

Noting that

$$
\begin{aligned}
(1-2\mu s)^{-\frac{1}{2}} &= \sum_{r=0}^{\infty} \frac{1}{r!} \frac{(2r)!}{r!} \left(\frac{1}{2}\right)^r \mu^r s^r, \qquad (4.104) \\
(1-2\mu s)^{-\frac{3}{2}} &= \sum_{r=0}^{\infty} \frac{1}{r!} \frac{(2r+2)!}{(r+1)!} \left(\frac{1}{2}\right)^{r+1} \mu^r s^r, \qquad (4.105)
\end{aligned}
$$

for small s, we may write (4.103) in the form

$$
\begin{aligned}
M_w(s) &= \sum_{r=0}^{\infty} \frac{1}{r!} \mu^r s^r \frac{(2r)!}{r!} \left(\frac{1}{2}\right)^r \\
&\quad + n^{-1} \left\{ b \sum_{r=1}^{\infty} \frac{1}{(r-1)!} s^r \mu^r \frac{(2r)!}{r!} \left(\frac{1}{2}\right)^r \right\} + O(n^{-\frac{3}{2}}). \qquad (4.106)
\end{aligned}
$$

By comparing (4.106) with (4.102) and noting that $\mu = -H^{-1}C$ here, we see that $w' = (1 + b/n)^{-1} w$ has a weighted χ_1^2-distribution with error of order $O(n^{-\frac{3}{2}})$ if and only if

$$b = (D + \delta^2)C^{-1} \tag{4.107}$$

and

$$F = G = 0. \tag{4.108}$$

This is an extremely restrictive condition, which in practice is likely only to be true when the model is correctly specified, so we consider now an alternative approach to obtaining a modified statistic.

4.1.5 Generalized Bartlett Adjustment when $\dim \psi = 1$

The form of (4.102) suggests that it may be possible to find a generalized Bartlett adjustment for w. Using (4.104), (4.105) and the analogous expansions

$$\begin{aligned}(1-2\mu s)^{-\frac{5}{2}} &= \frac{1}{3}\sum_{r=0}^{\infty}\frac{1}{r!}\frac{(2r+4)!}{(r+2)!}\left(\frac{1}{2}\right)^{r+2}\mu^r s^r \qquad (4.109)\\ (1-2\mu s)^{-\frac{7}{2}} &= \frac{1}{15}\sum_{r=0}^{\infty}\frac{1}{r!}\frac{(2r+6)!}{(r+3)!}\left(\frac{1}{2}\right)^{r+3}\mu^r s^r, \qquad (4.110)\end{aligned}$$

(4.102) may be written as

$$\begin{aligned} M_w(s) &= \sum_{r=0}^{\infty}\frac{1}{r!}\frac{(2r)!}{r!}\left(\frac{1}{2}\right)^r(-H^{-1}C)^r s^r \\ &\quad + n^{-1}\left\{\sum_{r=0}^{\infty}\frac{1}{(r+1)!}\frac{(2r+2)!}{r!}\left(\frac{1}{2}\right)^{r+1}(D+\delta^2)C^{-1}(-H^{-1}C)^{r+1}s^{r+1}\right. \\ &\quad + \sum_{r=0}^{\infty}\frac{1}{(r+2)!}\frac{(2r+4)!}{r!}\left(\frac{1}{2}\right)^r\left(\frac{1}{6}\delta F + \frac{1}{24}G\right)C^{-2}(-H^{-1}C)^{r+2}s^{r+2} \\ &\quad \left. + \sum_{r=0}^{\infty}\frac{1}{(r+3)!}\frac{(2r+6)!}{r!}\left(\frac{1}{2}\right)^r\frac{1}{72}F^2C^{-3}(-H^{-1}C)^{r+3}s^{r+3}\right\} + O(n^{-\frac{3}{2}}) \\ &= (1+2H^{-1}Cs)^{-\frac{1}{2}} \\ &\quad + n^{-1}\left\{\frac{1}{2}(D+\delta^2)C^{-1}\frac{-2H^{-1}Cs}{(1+2H^{-1}Cs)}\right. \\ &\qquad + \left(\frac{1}{2}\delta F + \frac{1}{8}G\right)C^{-2}\left[\frac{-2H^{-1}Cs}{(1+2H^{-1}Cs)}\right]^2 \\ &\qquad \left. + \frac{5}{24}F^2C^{-3}\left[\frac{-2H^{-1}Cs}{(1+2H^{-1}Cs)}\right]^3\right\} \\ &\quad + O(n^{-\frac{3}{2}}). \end{aligned} \tag{4.111}$$

Hence $-H^{-1}Cw$ has moment generating function

$$M_{-H^{-1}Cw}(s) = (1-2s)^{-\frac{1}{2}}\left\{1 + \frac{1}{24n}\left[A_1 d + A_2 d^2 + A_3 d^3\right]\right\}$$

$$+O(n^{-\frac{3}{2}}), \tag{4.112}$$

where

$$A_1 = 12\left(D+\delta^2\right)C^{-1} \tag{4.113}$$
$$A_2 = 12\left(\delta F+\frac{1}{4}G\right)C^{-2} \tag{4.114}$$
$$A_3 = 5F^2C^{-3}. \tag{4.115}$$

From proposition 2.1 it follows that if we define a statistic u by

$$u = -H^{-1}Cw \tag{4.116}$$

and put

$$u' = u\left\{1-(\gamma+\beta u+\alpha u^2)\right\}, \tag{4.117}$$

where

$$\alpha = -\frac{1}{36n}F^2C^{-3} \tag{4.118}$$
$$\beta = \frac{1}{n}\left\{\frac{5}{18}F^2C^{-3}-\frac{1}{3}\left(\delta F+\frac{1}{4}G\right)C^{-2}\right\} \tag{4.119}$$
$$\gamma = \frac{1}{n}\left\{\left(\delta F+\frac{1}{4}G\right)C^{-2}-\left(D+\delta^2\right)C^{-1}-\frac{5}{12}F^2C^{-3}\right\}, \tag{4.120}$$

then u' has a χ_1^2-distribution with error of order $O(n^{-\frac{3}{2}})$. Thus generalized Bartlett adjustment of a multiple of the likelihood ratio statistic under a mis-specified model is possible when the interest parameter, ψ, is a scalar.

4.2 The Score Test Statistic

Working in the ϕ-parameterization, we can obtain an expansion for the score test statistic, S_R, under the mis-specified model by Taylor expansion in the same way as for the null case. If, as before, we define the partitioned matrices

$$K = \begin{bmatrix} \nu_{\tau,\tau} & \nu_{\tau,\xi} \\ \nu_{\xi,\tau} & \nu_{\xi,\xi} \end{bmatrix}, \tag{4.121}$$
$$H = \begin{bmatrix} \nu_{\tau\tau} & \nu_{\tau\xi} \\ \nu_{\xi\tau} & \nu_{\xi\xi} \end{bmatrix}, \tag{4.122}$$

and from these the matrices

$$A = \begin{bmatrix} 0 & 0 \\ 0 & K_{\xi,\xi}^{-1} \end{bmatrix} \tag{4.123}$$
$$B = \begin{bmatrix} 0 & 0 \\ 0 & H_{\xi\xi}^{-1} \end{bmatrix}, \tag{4.124}$$

then, writing

$$S_R = S_0 + n^{-\frac{1}{2}} S_1 + n^{-1} S_2 + O\left(n^{-\frac{3}{2}}\right), \tag{4.125}$$

we find

$$\begin{aligned}
S_0 &= \nu^{h_1,h_2} Z_{h_1} Z_{h_2} && (4.126)\\
S_1 &= \nu^{h_1,\phi_1} \nu_{\xi_1,\phi_1,\phi_2} b^{\xi_1\xi_2} \nu^{\phi_2,h_2} Z_{\xi_2} Z_{h_1} Z_{h_2}\\
&\quad -2\nu^{h_1,h_2} b^{\xi_1\xi_2} Z_{h_1\xi_1} Z_{\xi_2} Z_{h_2}\\
&\quad +\nu^{h_1,h_2} b^{\xi_1\xi_3} b^{\xi_2\xi_4} \nu_{h_1\xi_1\xi_2} Z_{\xi_3} Z_{\xi_4} Z_{h_2} && (4.127)\\
S_2 &= -\nu^{h_1,\phi_1} \nu_{\xi_1,\phi_1,\phi_2} b^{\xi_1\xi_2} b^{\xi_3\xi_4} \nu^{\phi_2,h_2} Z_{h_1} Z_{h_2} Z_{\xi_2\xi_3} Z_{\xi_4}\\
&\quad +\frac{1}{2}\nu^{h_1,\phi_1} \nu_{\xi_1,\phi_1,\phi_2} b^{\xi_1\xi_2} \nu_{\xi_2\xi_3\xi_4} b^{\xi_3\xi_5} b^{\xi_4\xi_6} \nu^{\phi_2,h_2} Z_{\xi_6} Z_{\xi_5} Z_{h_1} Z_{h_2}\\
&\quad +\frac{1}{2}\nu^{h_1,\phi_1} \left(\nu_{\phi_1,\phi_2,\xi_1\xi_2} + 2\nu_{\phi_1,\xi_1,\phi_2\xi_2} + \nu_{\phi_1,\phi_2,\xi_1,\xi_2}\right) b^{\xi_1\xi_3} b^{\xi_2\xi_4} \nu^{\phi_2,h_2}\\
&\qquad\qquad Z_{\xi_3} Z_{\xi_4} Z_{h_1} Z_{h_2}\\
&\quad +\nu^{h_1,\phi_1} \nu_{\xi_1,\phi_1,\phi_2} b^{\xi_1\xi_2} \nu^{\phi_2\phi_3} \nu_{\xi_3,\phi_3,\phi_4} b^{\xi_3\xi_4} \nu^{\phi_4,h_2} Z_{\xi_2} Z_{\xi_4} Z_{h_1} Z_{h_2}\\
&\quad +2\nu^{h_1,h_2} b^{\xi_1\xi_2} b^{\xi_3\xi_4} Z_{h_1} Z_{h_2\xi_1} Z_{\xi_2\xi_3} Z_{\xi_4}\\
&\quad -2\nu^{h_1,h_2} \nu_{h_2\xi_1\xi_2} b^{\xi_1\xi_3} b^{\xi_2\xi_4} b^{\xi_5\xi_6} Z_{h_1} Z_{\xi_3} Z_{\xi_4\xi_5} Z_{\xi_6}\\
&\quad +\nu^{h_1,h_2} b^{\xi_1\xi_2} b^{\xi_3\xi_4} Z_{h_1} Z_{h_2\xi_1\xi_3} Z_{\xi_2} Z_{\xi_4}\\
&\quad -\frac{1}{3}\nu^{h_1,h_2} \nu_{h_2\xi_1\xi_2\xi_3} b^{\xi_1\xi_4} b^{\xi_2\xi_5} b^{\xi_3\xi_6} Z_{h_1} Z_{\xi_4} Z_{\xi_5} Z_{\xi_6}\\
&\quad +\nu^{h_1,h_2} \nu_{h_2\xi_1\xi_2} b^{\xi_1\xi_3} b^{\xi_2\xi_4} \nu_{\xi_4\xi_5\xi_6} b^{\xi_5\xi_7} b^{\xi_6\xi_8} Z_{h_1} Z_{\xi_3} Z_{\xi_7} Z_{\xi_8}\\
&\quad -\nu^{h_1,h_2} b^{\xi_1\xi_2} \nu_{\xi_2\xi_3\xi_4} b^{\xi_3\xi_5} b^{\xi_4\xi_6} Z_{h_1} Z_{h_2\xi_1} Z_{\xi_5} Z_{\xi_6}\\
&\quad +\nu^{h_1,h_2} b^{\xi_1\xi_2} b^{\xi_3\xi_4} Z_{h_1\xi_1} Z_{\xi_2} Z_{h_2\xi_3} Z_{\xi_4}\\
&\quad -\nu^{h_1,h_2} b^{\xi_1\xi_2} \nu_{h_2\xi_3\xi_4} b^{\xi_3\xi_5} b^{\xi_4\xi_6} Z_{h_1\xi_1} Z_{\xi_2} Z_{\xi_5} Z_{\xi_6}\\
&\quad +\frac{1}{4}\nu^{h_1,h_2} \nu_{h_1\xi_1\xi_2} b^{\xi_1\xi_3} b^{\xi_2\xi_4} \nu_{h_2\xi_5\xi_6} b^{\xi_5\xi_7} b^{\xi_6\xi_8} Z_{\xi_3} Z_{\xi_4} Z_{\xi_7} Z_{\xi_8}\\
&\quad -2\nu^{h_1,\phi_1} \nu_{\xi_1,\phi_1,\phi_2} b^{\xi_1\xi_2} \nu^{\phi_2 h_2} b^{\xi_3\xi_4} Z_{\xi_2} Z_{h_1} Z_{h_2\xi_3} Z_{\xi_4}\\
&\quad +\nu^{h_1,\phi_1} \nu_{\xi_1,\phi_1,\phi_2} b^{\xi_1\xi_2} \nu^{\phi_2,h_2} \nu_{h_2\xi_3\xi_4} b^{\xi_3\xi_5} b^{\xi_4\xi_6} Z_{\xi_2} Z_{h_1} Z_{\xi_5} Z_{\xi_6}. && (4.128)
\end{aligned}$$

It is easy to see that this expansion is exactly analogous to the null expansion of S_R, but the joint cumulants of the log likelihood here are those taken with respect to the true underlying density $g(x)$ rather than the null log-likelihood cumulants. As noted previously, the log-likelihood cumulants do not satisfy the condition

$$\nu_{r,s} = -\nu_{rs}, \tag{4.129}$$

so the simplifications used to calculate the moment generating function of the null score statistic cannot be used here and we must find some alternative.

4.2.1 Conditions for Generalized Bartlett Adjustment

The remainder of this section is concerned with the derivation of a generalized Bartlett adjustment of the score statistic under a mis-specified model, using a different approach to that used for the derivation under the null model. Consider the condition (3.7) given in section 3.1 for the existence of a generalized Bartlett adjustment. Rewriting (4.125) as

$$S_R = S_0(Z) + n^{-\frac{1}{2}} S_1(Z) + n^{-1} S_2(Z) + O\left(n^{-\frac{3}{2}}\right), \tag{4.130}$$

and expressing the true density function of $Z = \left(Z_\phi^T, Z_{\phi\phi}^T, Z_{\phi\phi\phi}^T\right)^T$ as $h(z)$, we have

$$\begin{aligned} M_{S_R}(t) &= \int e^{S_R t} g(x)\, dx \\ &= \int e^{\left\{S_0(z)+n^{-\frac{1}{2}}S_1(z)+n^{-1}S_2(z)\right\}t} h(z)\, dz + O\left(n^{-\frac{3}{2}}\right) \\ &= \int e^{S_0(z)t} h(z)\, dz + n^{-\frac{1}{2}} \int S_1(z) t e^{S_0(z)t} h(z)\, dz \\ &\quad + n^{-1} \int \left\{S_2(z) + \frac{1}{2}[S_1(z)]^2 t\right\} t e^{S_0(z)t} h(z)\, dz + O\left(n^{-\frac{3}{2}}\right). \end{aligned} \tag{4.131}$$

Noting that $S_1(Z)$ is an odd function of Z and $S_0(Z)$ is an even function of Z, so that

$$\begin{aligned} & \left[\int S_1(z) e^{S_0(z)t} h(z)\, dz\right]_{O(1)} \\ = & \left[\int e^{S_0(z)t} h(z)\, dz\right]_{O\left(n^{-\frac{1}{2}}\right)} \\ = & \; 0, \end{aligned} \tag{4.132}$$

we see that

$$\begin{aligned} \frac{M_1(t)}{t} &= \left[\frac{1}{t} \int e^{S_0(z)t} h(z)\, dz\right]_{O(n^{-1})} \\ &\quad + \left[\int S_1(z) e^{S_0(z)t} h(z)\, dz\right]_{O\left(n^{-\frac{1}{2}}\right)} \\ &\quad + \left[\int \left\{S_2(z) + \frac{1}{2}[S_1(z)]^2 t\right\} e^{S_0(z)t} h(z)\, dz\right]_{O(1)}. \end{aligned} \tag{4.133}$$

Now

$$\begin{aligned} \frac{d^r}{dt^r}\left\{\int e^{S_0(z)t}h(z)\,dz\right\} &= \int \frac{d^r}{dt^r}\left\{e^{S_0(z)t}h(z)\right\}\,dz \\ &= \int [S_0(z)]^r e^{S_0(z)t}h(z)\,dz \end{aligned} \tag{4.134}$$

so

$$\int e^{S_0(z)t}h(z)\,dz = \sum_{k=0}^{\infty}\frac{1}{k!}\mathrm{E}\left\{[S_0(Z)]^k\right\}t^k. \tag{4.135}$$

Finding expansions for the remaining terms of the expression (4.133) in the same way we have

$$\int S_1(z)e^{S_0(z)t}h(z)\,dz = \sum_{k=0}^{\infty}\frac{1}{k!}\mathrm{E}\left\{S_1(Z)[S_0(Z)]^k\right\}t^k \tag{4.136}$$

$$\int S_2(z)e^{S_0(z)t}h(z)\,dz = \sum_{k=0}^{\infty}\frac{1}{k!}\mathrm{E}\left\{S_2(Z)[S_0(Z)]^k\right\}t^k \tag{4.137}$$

$$t\int \frac{1}{2}[S_1(z)]^2 e^{S_0(z)t}h(z)\,dz = \frac{1}{2}\sum_{k=1}^{\infty}\frac{1}{(k-1)!}\mathrm{E}\left\{[S_1(Z)]^2[S_0(Z)]^{k-1}\right\}t^k \tag{4.138}$$

so

$$\begin{aligned} \frac{M_1(t)}{t} &= \left[\mathrm{E}\left\{S_1(Z)\right\}\right]_{O\left(n^{-\frac{1}{2}}\right)} + \left[\mathrm{E}\left\{S_2(Z)\right\}\right]_{O(1)} \\ &\quad + \sum_{k=1}^{\infty}\frac{1}{k!}t^k\left\{\left[\frac{1}{k+1}\mathrm{E}\left\{[S_0(Z)]^{k+1}\right\}\right]_{O(n^{-1})} + \left[\mathrm{E}\left\{S_1(Z)[S_0(Z)]^k\right\}\right]_{O\left(n^{-\frac{1}{2}}\right)}\right. \\ &\quad \left. + \left[\mathrm{E}\left\{S_2(Z)[S_0(Z)]^k + \frac{k}{2}[S_1(Z)]^2[S_0(Z)]^{k-1}\right\}\right]_{O(1)}\right\}. \end{aligned} \tag{4.139}$$

Now

$$\begin{aligned} M_0(t) &= \left[\int e^{S_0(z)t}h(z)\,dz\right]_{O(1)} \\ &= \sum_{k=0}^{\infty}\frac{1}{k!}t^k\left[\mathrm{E}\left\{[S_0(Z)]^k\right\}\right]_{O(1)}, \end{aligned} \tag{4.140}$$

so

$$\begin{aligned} M_0^{(r)}(t) &= \sum_{k=r}^{\infty}\frac{1}{(k-r)!}t^{k-r}\left[\mathrm{E}\left\{[S_0(Z)]^k\right\}\right]_{O(1)} \\ &= \sum_{k=0}^{\infty}\frac{1}{k!}t^k\left[\mathrm{E}\left\{[S_0(Z)]^{k+r}\right\}\right]_{O(1)}. \end{aligned} \tag{4.141}$$

By substituting (4.141) and (4.139) into (3.7), the condition for generalized Bartlett adjustment for the score test statistic may be written as

$$\sum_{r=0}^{d}a_r\sum_{k=0}^{\infty}\frac{1}{k!}t^k\left[\mathrm{E}\left\{[S_0(Z)]^{k+r+1}\right\}\right]_{O(1)} = -\frac{M_1(t)}{t}, \tag{4.142}$$

so

$$\begin{aligned}
& \sum_{k=0}^{\infty}\frac{1}{k!}t^k\left\{\sum_{r=0}^{d}a_r\Big[\mathrm{E}\left\{[S_0(Z)]^{k+r+1}\right\}\Big]_{O(1)}\right\} \\
= & -[\mathrm{E}\left\{S_1(Z)\right\}]_{O\left(n^{-\frac{1}{2}}\right)}-[\mathrm{E}\left\{S_2(Z)\right\}]_{O(1)} \\
& -\sum_{k=1}^{\infty}\frac{1}{k!}t^k\left\{\left[\frac{1}{k+1}\mathrm{E}\left\{[S_0(Z)]^{k+1}\right\}\right]_{O(n^{-1})}+\left[\mathrm{E}\left\{S_1(Z)[S_0(Z)]^k\right\}\right]_{O\left(n^{-\frac{1}{2}}\right)}\right. \\
& \left.+\left[\mathrm{E}\left\{S_2(Z)[S_0(Z)]^k+\frac{k}{2}[S_1(Z)]^2[S_0(Z)]^{k-1}\right\}\right]_{O(1)}\right\}. \qquad (4.143)
\end{aligned}$$

Thus if there are constants $a_1,\ldots,a_d$ which satisfy (4.143) then generalized Bartlett adjustment may be used to derive a modified score statistic of the form (3.2) under a mis-specified model.

4.2.2 Generalized Bartlett Adjustment when $\dim\tau=1$

For simplicity, consider the score statistic where the interest parameter, τ, is a scalar. In this case, the expressions (4.126) - (4.128) may be rewritten as

$$\begin{aligned}
S_0 &= \nu^{h,h}Z_h^2 \qquad (4.144) \\
S_1 &= \nu^{h,\phi_1}\nu_{\xi_1,\phi_1,\phi_2}b^{\xi_1\xi_2}\nu^{\phi_2,h}Z_{\xi_2}Z_h^2 \\
&\quad -2\nu^{h,h}b^{\xi_1\xi_2}Z_{h\xi_1}Z_{\xi_2}Z_h \\
&\quad +\nu^{h,h}b^{\xi_1\xi_3}b^{\xi_2\xi_4}\nu_{h\xi_1\xi_2}Z_{\xi_3}Z_{\xi_4}Z_h \qquad (4.145) \\
S_2 &= -\nu^{h,\phi_1}\nu_{\xi_1,\phi_1,\phi_2}b^{\xi_1\xi_2}b^{\xi_3\xi_4}\nu^{\phi_2,h}Z_h^2Z_{\xi_2\xi_3}Z_{\xi_4} \\
&\quad +\frac{1}{2}\nu^{h,\phi_1}\nu_{\xi_1,\phi_1,\phi_2}b^{\xi_1\xi_2}\nu_{\xi_2\xi_3\xi_4}b^{\xi_3\xi_5}b^{\xi_4\xi_6}\nu^{\phi_2,h}Z_{\xi_6}Z_{\xi_5}Z_h^2 \\
&\quad +\frac{1}{2}\nu^{h,\phi_1}\left(\nu_{\phi_1,\phi_2,\xi_1\xi_2}+2\nu_{\phi_1,\xi_1,\phi_2\xi_2}+\nu_{\phi_1,\phi_2,\xi_1,\xi_2}\right)b^{\xi_1\xi_3}b^{\xi_2\xi_4}\nu^{\phi_2,h} \\
&\qquad\quad Z_{\xi_3}Z_{\xi_4}Z_h^2 \\
&\quad +\nu^{h,\phi_1}\nu_{\xi_1,\phi_1,\phi_2}b^{\xi_1\xi_2}\nu^{\phi_2\phi_3}\nu_{\xi_3,\phi_3,\phi_4}b^{\xi_3\xi_4}\nu^{\phi_4,h}Z_{\xi_2}Z_{\xi_4}Z_h^2 \\
&\quad +2\nu^{h,h}b^{\xi_1\xi_2}b^{\xi_3\xi_4}Z_hZ_{h\xi_1}Z_{\xi_2\xi_3}Z_{\xi_4} \\
&\quad -2\nu^{h,h}\nu_{h\xi_1\xi_2}b^{\xi_1\xi_3}b^{\xi_2\xi_4}b^{\xi_5\xi_6}Z_hZ_{\xi_3}Z_{\xi_4\xi_5}Z_{\xi_6} \\
&\quad +\nu^{h,h}b^{\xi_1\xi_2}b^{\xi_3\xi_4}Z_hZ_{h\xi_1\xi_3}Z_{\xi_2}Z_{\xi_4} \\
&\quad -\frac{1}{3}\nu^{h,h}\nu_{h\xi_1\xi_2\xi_3}b^{\xi_1\xi_4}b^{\xi_2\xi_5}b^{\xi_3\xi_6}Z_hZ_{\xi_4}Z_{\xi_5}Z_{\xi_6} \\
&\quad +\nu^{h,h}\nu_{h\xi_1\xi_2}b^{\xi_1\xi_3}b^{\xi_2\xi_4}\nu_{\xi_4\xi_5\xi_6}b^{\xi_5\xi_7}b^{\xi_6\xi_8}Z_hZ_{\xi_3}Z_{\xi_7}Z_{\xi_8} \\
&\quad -\nu^{h,h}b^{\xi_1\xi_2}\nu_{\xi_2\xi_3\xi_4}b^{\xi_3\xi_5}b^{\xi_4\xi_6}Z_hZ_{h\xi_1}Z_{\xi_5}Z_{\xi_6}
\end{aligned}$$

$$
\begin{aligned}
&+\nu^{h,h} b^{\xi_1\xi_2} b^{\xi_3\xi_4} Z_{h\xi_1} Z_{\xi_2} Z_{h\xi_3} Z_{\xi_4} \\
&-\nu^{h,h} b^{\xi_1\xi_2} \nu_{h\xi_3\xi_4} b^{\xi_3\xi_5} b^{\xi_4\xi_6} Z_{h\xi_1} Z_{\xi_2} Z_{\xi_5} Z_{\xi_6} \\
&+\frac{1}{4}\nu^{h,h} \nu_{h\xi_1\xi_2} b^{\xi_1\xi_3} b^{\xi_2\xi_4} \nu_{h\xi_5\xi_6} b^{\xi_5\xi_7} b^{\xi_6\xi_8} Z_{\xi_3} Z_{\xi_4} Z_{\xi_7} Z_{\xi_8} \\
&-2\nu^{h,\phi_1} \nu_{\xi_1,\phi_1,\phi_2} b^{\xi_1\xi_2} \nu^{\phi_2 h} b^{\xi_3\xi_4} Z_{\xi_2} Z_h Z_{h\xi_3} Z_{\xi_4} \\
&+\nu^{h,\phi_1} \nu_{\xi_1,\phi_1,\phi_2} b^{\xi_1\xi_2} \nu^{\phi_2,h} \nu_{h\xi_3\xi_4} b^{\xi_3\xi_5} b^{\xi_4\xi_6} Z_{\xi_2} Z_h Z_{\xi_5} Z_{\xi_6}. \qquad (4.146)
\end{aligned}
$$

From the definitions of the joint cumulants of log-likelihood derivatives we have

$$
\begin{aligned}
\mathrm{E}\left[Z_r Z_s\right] &= \nu_{r,s}, & (4.147)\\
\mathrm{E}\left[Z_r Z_s Z_t\right] &= n^{-\frac{1}{2}} \nu_{r,s,t}, & (4.148)\\
\mathrm{E}\left[Z_r Z_s Z_t Z_u\right] &= n^{-1} \nu_{r,s,t,u} + \left(\nu_{r,s}\nu_{t,u}[3]\right), & (4.149)
\end{aligned}
$$

and so on. It follows that

$$
\left[\mathrm{E}\left\{Z_{i_1} \ldots Z_{i_{2r}}\right\}\right]_{O(1)} = \prod_{j=1}^{r} \nu_{i_{2j},i_{2j-1}} \left[\frac{(2r)!}{2^r r!}\right], \qquad (4.150)
$$

$$
\left[\mathrm{E}\left\{Z_{i_1} \ldots Z_{i_{2r+1}}\right\}\right]_{O\left(n^{-\frac{1}{2}}\right)} = \nu_{i_1,i_2,i_3} \prod_{j=2}^{r} \nu_{i_{2j},i_{2j+1}} \left[\frac{(2r+1)!}{3.2^r.(r-1)!}\right], \qquad (4.151)
$$

$$
\left[\mathrm{E}\left\{Z_{i_1} \ldots Z_{i_{2r}}\right\}\right]_{O(n^{-1})} = \begin{cases} 0 & r=1 \\ \nu_{i_1,i_2,i_3,i_4} & r=2 \\ \nu_{i_1,i_2,i_3,i_4} \prod_{j=3}^{r} \nu_{i_{2j},i_{2j+1}} \left[\frac{(2r)!}{4!2^{r-2}(r-2)!}\right] & \\ +\nu_{i_1,i_2,i_3}\nu_{i_4,i_5,i_6} & \\ \qquad \prod_{j=4}^{r} \nu_{i_{2j},i_{2j+1}} \left[\frac{(2r)!}{(3!)^2 2^{r-2}(r-3)!}\right] & r \geq 3. \end{cases} \qquad (4.152)
$$

Using (4.150) - (4.152) we find that, for $r \geq 2$,

$$
\begin{aligned}
& \mathrm{E}\left\{\left[S_0(Z)\right]^{r+1}\right\} \\
=\ & \mathrm{E}\left\{\left(\nu^{h,h}\right)^{r+1} Z_h^{2r+2}\right\} \\
=\ & \left(\nu^{h,h}\right)^{r+1} \left(\nu_{h,h}\right)^{r+1} \frac{(2r+2)!}{2^{r+1}(r+1)!} \\
& +n^{-1}\left\{\left(\nu^{h,h}\right)^{r+1} \nu_{h,h,h,h} \left(\nu_{h,h}\right)^{r-1} \frac{(2r+2)!}{4!2^{r-1}(r-1)!}\right. \\
& \qquad \left. +\left(\nu^{h,h}\right)^{r+1} \left(\nu_{h,h,h}\right)^2 \left(\nu_{h,h}\right)^{r-2} \frac{(2r+2)!}{(3!)^2 2^{r-1}(r-2)!}\right\} \\
& +O\left(n^{-\frac{3}{2}}\right) \qquad (4.153)
\end{aligned}
$$

$$
\begin{aligned}
& \left[\mathrm{E}\left\{S_1(Z)[S_0(Z)]^r\right\}\right]_{O\left(n^{-\frac{1}{2}}\right)} \\
= \;& \Big[\mathrm{E}\Big\{\nu^{h,\phi_1}\nu_{\xi_1,\phi_1,\phi_2}b^{\xi_1\xi_2}\nu^{\phi_2,h}Z_{\xi_2}\left(\nu^{h,h}\right)^r Z_h^{2r+2} \\
& \quad - 2\nu^{h,h}b^{\xi_1\xi_2}Z_{h\xi_1}Z_{\xi_2}\left(\nu^{h,h}\right)^r Z_h^{2r+1} \\
& \quad + \nu^{h,h}b^{\xi_1\xi_3}b^{\xi_2\xi_4}\nu_{h\xi_1\xi_2}Z_{\xi_3}Z_{\xi_4}\left(\nu^{h,h}\right)^r Z_h^{2r+1}\Big\}\Big]_{O\left(n^{-\frac{1}{2}}\right)} \\
= \;& \nu^{h,\phi_1}\nu_{\xi_1,\phi_1,\phi_2}b^{\xi_1\xi_2}\nu^{\phi_2,h}\left(\nu^{h,h}\right)^r \\
& \qquad \left\{\nu_{\xi_2,h,h}(\nu_{h,h})^r\frac{(2r+2)!}{2^{r+1}r!} + \nu_{h,h,h}\nu_{\xi_2,h}(\nu_{h,h})^{r-1}\frac{(2r+2)!}{3.2^r(r-1)!}\right\} \\
& -2\left(\nu^{h,h}\right)^{r+1}b^{\xi_1\xi_2}\left\{\nu_{h\xi_1,\xi_2,h}(\nu_{h,h})^r\frac{(2r+1)!}{2^r r!} + \nu_{h\xi_1,h,h}\nu_{\xi_2,h}(\nu_{h,h})^{r-1}\frac{(2r+1)!}{2^r(r-1)!}\right\} \\
& +\left(\nu^{h,h}\right)^{r+1}b^{\xi_1\xi_3}b^{\xi_2\xi_4}\nu_{h\xi_1\xi_2} \\
& \qquad \Big\{\nu_{\xi_3,\xi_4}\nu_{h,h,h}(\nu_{h,h})^{r-1}\frac{(2r+1)!}{3.2^r(r-1)!} \\
& \qquad + \nu_{\xi_3,h}\nu_{\xi_4,h}\nu_{h,h,h}(\nu_{h,h})^{r-2}\frac{(2r+1)!}{3.2^{r-1}(r-2)!} \\
& \qquad + \nu_{\xi_3,h,h}\nu_{\xi_4,h}(\nu_{h,h})^{r-1}\frac{(2r+1)!}{2^{r+1}(r-1)!} \\
& \qquad + \nu_{\xi_3,\xi_4,h}(\nu_{h,h})^r\frac{(2r+1)!}{2^r r!}\Big\},
\end{aligned}
\tag{4.154}
$$

$$
\begin{aligned}
& \left[\mathrm{E}\left\{S_2(Z)[S_0(Z)]^r\right\}\right]_{O(1)} \\
= \;& \frac{1}{2}\nu^{h,\phi_1}\nu_{\xi_1,\phi_1,\phi_2}b^{\xi_1\xi_2}\nu_{\xi_2\xi_3\xi_4}b^{\xi_3\xi_5}b^{\xi_4\xi_6}\nu^{\phi_2,h}\left(\nu^{h,h}\right)^r \\
& \qquad \Big\{\nu_{\xi_5,\xi_6}(\nu_{h,h})^{r+1}\frac{(2r+2)!}{2^{r+1}(r+1)!} \\
& \qquad + \nu_{\xi_5,h}\nu_{\xi_6,h}(\nu_{h,h})^r\frac{(2r+2)!}{2^r r!}\Big\} \\
& + \frac{1}{2}\nu^{h,\phi_1}\left(\nu_{\phi_1,\phi_2,\xi_1\xi_2} + 2\nu_{\phi_1,\xi_1,\phi_2\xi_2} + \nu_{\phi_1,\phi_2,\xi_1,\xi_2}\right)\left(\nu^{h,h}\right)^r b^{\xi_1\xi_3}b^{\xi_2\xi_4}\nu^{\phi_2,h} \\
& \qquad \Big\{\nu_{\xi_3,\xi_4}(\nu_{h,h})^{r+1}\frac{(2r+2)!}{2^{r+1}(r+1)!} \\
& \qquad + \nu_{\xi_2,h}\nu_{\xi_4,h}(\nu_{h,h})^r\frac{(2r+2)!}{2^r r!}\Big\} \\
& + \nu^{h,\phi_1}\nu_{\xi_1,\phi_1,\phi_2}b^{\xi_1\xi_2}\nu^{\phi_2,\phi_3}\nu_{\xi_3,\phi_3,\phi_4}b^{\xi_3\xi_4}\nu^{\phi_4,h}\left(\nu^{h,h}\right)^r \\
& \qquad \Big\{\nu_{\xi_2,\xi_4}(\nu_{h,h})^{r+1}\frac{(2r+2)!}{2^{r+1}(r+1)!}
\end{aligned}
$$

$$
\begin{aligned}
& +\nu_{\xi_2,h}\nu_{\xi_4,h}\left(\nu_{h,h}\right)^r\frac{(2r+2)!}{2^r r!}\Bigg\} \\
& +2b^{\xi_1\xi_2}b^{\xi_3\xi_4}\left(\nu^{h,h}\right)^{r+1}\nu_{h\xi_1,\xi_2\xi_3}\nu_{\xi_4,h}\left(\nu_{h,h}\right)^r\frac{(2r+1)!}{2^r r!} \\
& -\frac{1}{3}\nu_{h\xi_1\xi_2\xi_3}b^{\xi_1\xi_4}b^{\xi_2\xi_5}b^{\xi_3\xi_6}\left(\nu^{h,h}\right)^{r+1} \\
& \qquad \Bigg\{\left(\nu_{\xi_4,\xi_5}\nu_{\xi_6,h}[3]\right)\left(\nu_{h,h}\right)^r\frac{(2r+1)!}{2^r r!} \\
& \qquad +\nu_{\xi_4,h}\nu_{\xi_5,h}\nu_{\xi_6,h}\left(\nu_{h,h}\right)^{r-1}\frac{(2r+1)!}{2^{r-1}(r-1)!}\Bigg\} \\
& +\nu_{h\xi_1\xi_2}b^{\xi_1\xi_3}b^{\xi_2\xi_4}\nu_{\xi_4\xi_5\xi_6}b^{\xi_5\xi_7}b^{\xi_6\xi_8}\left(\nu^{h,h}\right)^{r+1} \\
& \qquad \Bigg\{\left(\nu_{\xi_7,\xi_8}\nu_{\xi_3,h}[3]\right)\left(\nu_{h,h}\right)^r\frac{(2r+1)!}{2^r r!} \\
& \qquad +\nu_{\xi_3,h}\nu_{\xi_7,h}\nu_{\xi_8,h}\left(\nu_{h,h}\right)^{r-1}\frac{(2r+1)!}{2^{r-1}(r-1)!}\Bigg\} \\
& +b^{\xi_1\xi_2}b^{\xi_3\xi_4}\left(\nu^{h,h}\right)^{r+1}\nu_{h\xi_1,h\xi_3} \\
& \qquad \Bigg\{\nu_{\xi_2,\xi_4}\left(\nu_{h,h}\right)^r\frac{(2r)!}{2^r r!} \\
& \qquad \nu_{\xi_2,h}\nu_{\xi_4,h}\left(\nu_{h,h}\right)^{r-1}\frac{(2r)!}{2^{r-1}(r-1)!}\Bigg\} \\
& +\frac{1}{4}\nu_{h\xi_1\xi_2}b^{\xi_1\xi_3}b^{\xi_2\xi_4}\nu_{h\xi_5\xi_6}b^{\xi_5\xi_7}b^{\xi_6\xi_8}\left(\nu^{h,h}\right)^{r+1} \\
& \qquad \Bigg\{\left(\nu_{\xi_3,\xi_4}\nu_{\xi_7,\xi_8}[3]\right)\left(\nu_{h,h}\right)^r\frac{(2r)!}{2^r r!} \\
& \qquad +\left(\nu_{\xi_3,\xi_4}\nu_{\xi_7,h}\nu_{\xi_8,h}[6]\right)\left(\nu_{h,h}\right)^{r-1}\frac{(2r)!}{2^{r-1}(r-1)!} \\
& \qquad +\nu_{\xi_3,h}\nu_{\xi_4,h}\nu_{\xi_7,h}\nu_{\xi_8,h}\left(\nu_{h,h}\right)^{r-2}\frac{(2r)!}{2^{r-2}(r-2)!}\Bigg\} \\
& \nu^{h,\phi_1}\nu_{\xi_1,\phi_1,\phi_2}b^{\xi_1\xi_2}\nu^{\phi_2,h}\nu_{h\xi_3\xi_4}b^{\xi_3\xi_5}b^{\xi_4\xi_6}\left(\nu^{h,h}\right)^r \\
& \qquad \Bigg\{\left(\nu_{\xi_5,\xi_6}\nu_{\xi_2,h}[3]\right)\left(\nu_{h,h}\right)^r\frac{(2r+1)!}{2^r r!} \\
& \qquad +\nu_{\xi_5,h}\nu_{\xi_6,h}\nu_{\xi_2,h}\left(\nu_{h,h}\right)^{r-1}\frac{(2r+1)!}{2^{r-1}(r-1)!}\Bigg\},
\end{aligned}
\tag{4.155}
$$

$$
\begin{aligned}
& \left[\mathrm{E}\left\{[S_1(Z)]^2[S_0(Z)]^{r-1}\right\}\right]_{O(1)} \\
= \ & \nu^{h,\phi_1}\nu_{\xi_1,\phi_1,\phi_2}b^{\xi_1\xi_2}\nu^{\phi_2,h}\nu^{h,\phi_3}\nu_{\xi_3,\phi_3,\phi_4}b^{\xi_3\xi_4}\nu^{\phi_4,h}\left(\nu^{h,h}\right)^{r-1} \\
& \qquad \Bigg\{\nu_{\xi_2,\xi_4}\left(\nu_{h,h}\right)^{r+1}\frac{(2r+2)!}{2^{r+1}(r+1)!}
\end{aligned}
$$

$$
\begin{aligned}
&\quad +\nu_{\xi_2,h}\nu_{\xi_4,h}(\nu_{h,h})^r \frac{(2r+2)!}{2^r r!}\bigg\} \\
&+4b^{\xi_1\xi_2}b^{\xi_3\xi_4}\left(\nu^{h,h}\right)^{r+1}\nu_{h\xi_1,h\xi_3} \\
&\quad \bigg\{\nu_{\xi_2,\xi_4}(\nu_{h,h})^r \frac{(2r)!}{2^r r!} \\
&\quad +\nu_{\xi_2,h}\nu_{\xi_4,h}(\nu_{h,h})^{r-1}\frac{(2r)!}{2^{r-1}(r-1)!}\bigg\} \\
&+b^{\xi_1\xi_3}b^{\xi_2\xi_4}b^{\xi_5\xi_7}b^{\xi_6\xi_8}\nu_{h\xi_1\xi_2}\nu_{h\xi_5\xi_6}\left(\nu^{h,h}\right)^{r+1} \\
&\quad \bigg\{\left(\nu_{\xi_3,\xi_4}\nu_{\xi_7,\xi_8}[3]\right)(\nu_{h,h})^r \frac{(2r)!}{2^r r!} \\
&\quad +\left(\nu_{\xi_3,\xi_4}\nu_{\xi_7,h}\nu_{\xi_8,h}[6]\right)(\nu_{h,h})^{r-1}\frac{(2r)!}{2^{r-1}(r-1)!} \\
&\quad +\nu_{\xi_3,h}\nu_{\xi_4,h}\nu_{\xi_7,h}\nu_{\xi_8,h}(\nu_{h,h})^{r-2}\frac{(2r)!}{2^{r-2}(r-2)!}\bigg\} \\
&+2\nu^{h,\phi_1}\nu_{\xi_1,\phi_1,\phi_2}b^{\xi_1\xi_2}\nu^{\phi_2,h}b^{\xi_3\xi_5}b^{\xi_4\xi_6}\nu_{h\xi_3\xi_4}\left(\nu^{h,h}\right)^r \\
&\quad \bigg\{\left(\nu_{\xi_5,\xi_6}\nu_{\xi_2,h}[3]\right)(\nu_{h,h})^r \frac{(2r+1)!}{2^r r!} \\
&\quad +\nu_{\xi_2,h}\nu_{\xi_5,h}\nu_{\xi_6,h}(\nu_{h,h})^{r-1}\frac{(2r+1)!}{2^{r-1}(r-1)!}\bigg\}. \qquad (4.156)
\end{aligned}
$$

On substituting the $O(1)$ term of (4.153) into the left-hand side of equation (4.143), we find that

$$
\begin{aligned}
& \sum_{k=0}^{\infty}\frac{1}{k!}\sum_{r=0}^{d} a_r\left[\mathrm{E}\left\{[S_0(Z)]^{k+r+1}\right\}\right]_{O(1)} t^k \\
= & \sum_{k=0}^{\infty}\frac{1}{k!}\sum_{r=0}^{d} a_r\frac{(2k+2r+2)!}{2^{k+r+1}(k+r+1)!}t^k \\
= & \sum_{k=0}^{\infty}\frac{1}{k!}\frac{(2k+2)!}{2^{k+1}(k+1)!} \\
& \quad \left\{a_0+\frac{(2k+3)(2k+4)}{2(k+2)}a_1+\ldots+\frac{(2k+3)\ldots(2k+2d+2)}{2^d(k+2)\ldots(k+d+1)}a_d\right\}t^k \\
= & \sum_{k=0}^{\infty}\frac{1}{k!}\frac{(2k+2)!}{2^{k+1}(k+1)!}\left\{\sum_{r=0}^{d}a_r\prod_{i=1}^{r}(2k+2i+1)\right\}t^k \qquad (4.157)
\end{aligned}
$$

and the coefficient of $\frac{1}{k!}t^k$ in the right-hand side of (4.143) for $k \geq 2$ is

$$
\begin{aligned}
&\left[\frac{1}{k+1}\mathrm{E}\left\{[S_0(Z)]^{k+1}\right\}\right]_{O(n^{-1})}+\left[\mathrm{E}\left\{S_1(Z)[S_0(Z)]^k\right\}\right]_{O\left(n^{-\frac{1}{2}}\right)} \\
&+\left[\mathrm{E}\left\{S_2(Z)[S_0(Z)]^k+\frac{k}{2}[S_1(Z)]^2[S_0(Z)]^{k-1}\right\}\right]_{O(1)}
\end{aligned}
$$

$$
\begin{aligned}
= \; & \frac{(2k+2)!}{2^{k+1}(k+1)!} \\
& \Big\{ \left(\nu^{h,h}\right)^2 \nu_{h,h,h,h} \left(\frac{1}{6}k\right) \\
& + \left(\nu^{h,h}\right)^3 (\nu_{h,h,h})^2 \left(\frac{1}{9}k^2 - \frac{1}{9}k\right) \\
& + \nu^{h,\phi_1} \nu_{\xi_1,\phi_1,\phi_2} b^{\xi_1\xi_2} \nu^{\phi_2,h} \\
& \qquad\qquad [\nu_{\xi_2,h,h}(k+1) \\
& \qquad\qquad + \nu_{h,h,h}\nu_{\xi_2,h}\nu^{h,h}\left(\frac{4}{3}k^2 + \frac{4}{3}k\right)\Big] \\
& + b^{\xi_1\xi_2}\Big[\nu_{h\xi_1,\xi_2,h}\nu^{h,h}(-2) \\
& \qquad + \nu_{h\xi_1,h,h}\nu_{\xi_2,h}\left(\nu^{h,h}\right)^2(-2k)\Big] \\
& + b^{\xi_1\xi_3}b^{\xi_2\xi_4}\nu_{h\xi_1\xi_2}\Big[\nu_{\xi_3,\xi_4}\nu_{h,h,h}\left(\nu^{h,h}\right)^2\left(\frac{1}{3}k\right) \\
& \qquad\qquad + \nu_{\xi_3,h}\nu_{\xi_4,h}\nu_{h,h,h}\left(\nu^{h,h}\right)^3\left(\frac{2}{3}k^2 - \frac{2}{3}k\right) \\
& \qquad\qquad + \nu_{\xi_3,h,h}\nu_{\xi_4,h}\left(\nu^{h,h}\right)^2\left(\frac{1}{2}k\right) \\
& \qquad\qquad + \nu_{\xi_3,\xi_4,h}\nu^{h,h}(1)\Big] \\
& + \nu^{h,\phi_1}\nu_{\xi_1,\phi_1,\phi_2}b^{\xi_1\xi_2}\nu_{\xi_2\xi_3\xi_4}b^{\xi_3\xi_5}b^{\xi_4\xi_6}\nu^{\phi_2,h}\Big[\nu_{\xi_5,\xi_6}\nu_{h,h}\left(\frac{1}{2}\right) \\
& \qquad\qquad\qquad + \nu_{\xi_5,h}\nu_{\xi_6,h}(k+1)\Big] \\
& + \nu^{h,\phi_1}\left(\nu_{\phi_1,\phi_2,\xi_1\xi_2} + 2\nu_{\phi_1,\xi_1,\phi_2\xi_2} + \nu_{\phi_1,\phi_2,\xi_1,\xi_2}\right)b^{\xi_1\xi_3}b^{\xi_2\xi_4}\nu^{\phi_2,h} \\
& \qquad\qquad \Big[\nu_{\xi_3,\xi_4}\nu_{h,h}\left(\frac{1}{2}\right) \\
& \qquad\qquad + \nu_{\xi_3,h}\nu_{\xi_4,h}(k+1)] \\
& + \nu^{h,\phi_1}\nu_{\xi_1,\phi_1,\phi_2}b^{\xi_1\xi_2}\nu^{\phi_2,\phi_3}\nu_{\xi_3,\phi_3,\phi_4}b^{\xi_3\xi_4}\nu^{\phi_4,h} \\
& \qquad\qquad [\nu_{\xi_2,\xi_4}\nu_{h,h}(1) \\
& \qquad\qquad + \nu_{\xi_2,h}\nu_{\xi_4,h}(2k+2)] \\
& + b^{\xi_1\xi_2}b^{\xi_3\xi_4}\nu_{h\xi_1,\xi_2\xi_3}\nu_{\xi_4,h}\nu^{h,h}(2) \\
& + \nu_{h\xi_1\xi_2\xi_3}b^{\xi_1\xi_4}b^{\xi_2\xi_5}b^{\xi_3\xi_6} \\
& \qquad\qquad \Big[\left(\nu_{\xi_4,\xi_5}\nu_{\xi_6,h}[3]\right)\nu^{h,h}\left(-\frac{1}{3}\right) \\
& \qquad\qquad + \nu_{\xi_4,h}\nu_{\xi_5,h}\nu_{\xi_6,h}\left(\nu^{h,h}\right)^2\left(-\frac{2}{3}k\right)\Big] \\
& + \nu_{h\xi_1\xi_2}b^{\xi_1\xi_3}b^{\xi_2\xi_4}\nu_{\xi_4\xi_5\xi_6}b^{\xi_5\xi_7}b^{\xi_6\xi_8} \\
& \qquad\qquad \Big[\left(\nu_{\xi_7,\xi_8}\nu_{\xi_3,h}[3]\right)\nu^{h,h}(1)
\end{aligned}
$$

$$
\begin{aligned}
&\qquad +\nu_{\xi_3,h}\nu_{\xi_7,h}\nu_{\xi_8,h}\left(\nu^{h,h}\right)^2(2k)\Big] \\
&+ b^{\xi_1\xi_2}b^{\xi_3\xi_4}\nu_{h\xi_1,h\xi_3} \\
&\qquad \left[\nu_{\xi_2,\xi_4}\nu^{h,h}\left(\frac{1}{2k+1}\right)\right. \\
&\qquad \left. +\nu_{\xi_2,h}\nu_{\xi_4,h}\left(\nu^{h,h}\right)^2\left(\frac{2k}{2k+1}\right)\right] \\
&+ \nu_{h\xi_1\xi_2}b^{\xi_1\xi_3}b^{\xi_2\xi_4}\nu_{h\xi_5\xi_6}b^{\xi_5\xi_7}b^{\xi_6\xi_8} \\
&\qquad \left[\left(\nu_{\xi_3,\xi_4}\nu_{\xi_7,\xi_8}[3]\right)\nu^{h,h}\left(\frac{1}{8k+4}\right)\right. \\
&\qquad + \left(\nu_{\xi_3,\xi_4}\nu_{\xi_7,h}\nu_{\xi_8,h}[6]\right)\left(\nu^{h,h}\right)^2\left(\frac{k}{4k+2}\right) \\
&\qquad \left. +\nu_{\xi_3,h}\nu_{\xi_4,h}\nu_{\xi_7,h}\nu_{\xi_8,h}\left(\nu^{h,h}\right)^3\left(\frac{k(k-1)}{2k+1}\right)\right] \\
&+ \nu^{h,\phi_1}\nu_{\xi_1,\phi_1,\phi_2}b^{\xi_1\xi_2}\nu^{\phi_2,h}\nu_{h\xi_3\xi_4}b^{\xi_3\xi_5}b^{\xi_4\xi_6} \\
&\qquad \Big[\left(\nu_{\xi_5,\xi_6}\nu_{\xi_2,h}[3]\right)(1) \\
&\qquad +\nu_{\xi_5,h}\nu_{\xi_6,h}\nu_{\xi_2,h}\nu^{h,h}(2k)\Big] \\
&+ \nu^{h,\phi_1}\nu_{\xi_1,\phi_1,\phi_2}b^{\xi_1\xi_2}\nu^{\phi_2,h}\nu^{h,\phi_3}\nu_{\xi_3,\phi_3,\phi_4}b^{\xi_3\xi_4}\nu^{\phi_4,h} \\
&\qquad \left[\nu_{\xi_2,\xi_4}(\nu_{h,h})^2\left(\frac{k}{2}\right)\right. \\
&\qquad \left. + \nu_{\xi_2,h}\nu_{\xi_4,h}\nu_{h,h}\left(k^2+k\right)\right] \\
&+ b^{\xi_1\xi_2}b^{\xi_3\xi_4}\nu_{h\xi_1,h\xi_3} \\
&\qquad \left[\nu_{\xi_2,\xi_4}\nu^{h,h}\left(\frac{2k}{2k+1}\right)\right. \\
&\qquad \left. +\nu_{\xi_2,h}\nu_{\xi_4,h}\left(\nu^{h,h}\right)^2\left(\frac{4k^2}{2k+1}\right)\right] \\
&b^{\xi_1\xi_3}b^{\xi_2\xi_4}b^{\xi_5\xi_7}b^{\xi_6\xi_8}\nu_{h\xi_1\xi_2}\nu_{h\xi_5\xi_6} \\
&\qquad \left[\left(\nu_{\xi_3,\xi_4}\nu_{\xi_7,\xi_8}[3]\right)\nu^{h,h}\left(\frac{k}{4k+2}\right)\right. \\
&\qquad + \left(\nu_{\xi_3,\xi_4}\nu_{\xi_7,h}\nu_{\xi_8,h}[6]\right)\left(\nu^{h,h}\right)^2\left(\frac{k^2}{2k+1}\right) \\
&\qquad \left. +\nu_{\xi_3,h}\nu_{\xi_4,h}\nu_{\xi_7,h}\nu_{\xi_8,h}\left(\nu^{h,h}\right)^3\left(\frac{2k^3-2k^2}{2k+1}\right)\right] \\
&+ \nu^{h,\phi_1}\nu_{\xi_1,\phi_1,\phi_2}b^{\xi_1\xi_2}\nu^{\phi_2,h}b^{\xi_3\xi_5}b^{\xi_4\xi_6}\nu_{h\xi_3\xi_4} \\
&\qquad \Big[\left(\nu_{\xi_5,\xi_6}\nu_{\xi_2,h}[3]\right)(k) \\
&\qquad +\nu_{\xi_2,h}\nu_{\xi_5,h}\nu_{\xi_6,h}\nu^{h,h}\left(2k^2\right)\Big].
\end{aligned}
\tag{4.158}
$$

Gathering together powers of k we find that

$$\begin{aligned}
&\left[\frac{1}{k+1}\mathrm{E}\left\{[S_0(Z)]^{k+1}\right\}\right]_{O(n^{-1})}+\left[\mathrm{E}\left\{S_1(Z)[S_0(Z)]^k\right\}\right]_{O\left(n^{-\frac{1}{2}}\right)} \\
&+\left[\mathrm{E}\left\{S_2(Z)[S_0(Z)]^k+\frac{k}{2}[S_1(Z)]^2[S_0(Z)]^{k-1}\right\}\right]_{O(1)} \\
=\ & \frac{(2k+2)!}{2^{k+1}(k+1)!}\left[D_0+D_1k+D_2k^2\right], \qquad (4.159)
\end{aligned}$$

where

$$\begin{aligned}
D_0 \;=\;& \nu^{h,\phi_1}\nu_{\xi_1,\phi_1,\phi_2}b^{\xi_1\xi_2}\nu^{\phi_2,h}\nu_{\xi_2,h,h} \\
&-2b^{\xi_1\xi_2}\nu_{h\xi_1,\xi_2,h}\nu^{h,h} \\
&+b^{\xi_1\xi_3}b^{\xi_2\xi_4}\nu_{h\xi_1\xi_2}\nu_{\xi_3,\xi_4,h}\nu^{h,h} \\
&+\frac{1}{2}\nu^{h,\phi_1}\nu_{\xi_1,\phi_1,\phi_2}b^{\xi_1\xi_2}\nu_{\xi_2\xi_3\xi_4}b^{\xi_3\xi_5}b^{\xi_4\xi_6}\nu^{\phi_2,h}\nu_{\xi_5,\xi_6}\nu_{h,h} \\
&+\nu^{h,\phi_1}\nu_{\xi_1,\phi_1,\phi_2}b^{\xi_1\xi_2}\nu_{\xi_2\xi_3\xi_4}b^{\xi_3\xi_5}b^{\xi_4\xi_6}\nu^{\phi_2,h}\nu_{\xi_5,h}\nu_{\xi_6,h} \\
&+\frac{1}{2}\nu^{h,\phi_1}\left(\nu_{\phi_1,\phi_2,\xi_1\xi_2}+2\nu_{\phi_1,\xi_1,\phi_2\xi_2}+\nu_{\phi_1,\phi_2,\xi_1,\xi_2}\right)b^{\xi_1\xi_3}b^{\xi_2\xi_4}\nu^{\phi_2,h}\nu_{\xi_3,\xi_4}\nu_{h,h} \\
&+\nu^{h,\phi_1}\left(\nu_{\phi_1,\phi_2,\xi_1\xi_2}+2\nu_{\phi_1,\xi_1,\phi_2\xi_2}+\nu_{\phi_1,\phi_2,\xi_1,\xi_2}\right)b^{\xi_1\xi_3}b^{\xi_2\xi_4}\nu^{\phi_2,h}\nu_{\xi_3,h}\nu_{\xi_4,h} \\
&+\nu^{h,\phi_1}\nu_{\xi_1,\phi_1,\phi_2}b^{\xi_1\xi_2}\nu^{\phi_2,\phi_3}\nu_{\xi_2,\phi_3,\phi_4}b^{\xi_3,\xi_4}\nu^{\phi_4,h}\nu_{h,h}\nu_{\xi_3,\xi_4} \\
&+2\nu^{h,\phi_1}\nu_{\xi_1,\phi_1,\phi_2}b^{\xi_1\xi_2}\nu^{\phi_2,\phi_3}\nu_{\xi_2,\phi_3,\phi_4}b^{\xi_3,\xi_4}\nu^{\phi_4,h}\nu_{\xi_3,h}\nu_{\xi_4,h} \\
&+2b^{\xi_1\xi_2}b^{\xi_3\xi_4}\nu_{h\xi_1,\xi_2\xi_3}\nu_{\xi_4,h}\nu^{h,h} \\
&-\frac{1}{3}\nu_{h\xi_1\xi_2\xi_3}b^{\xi_1\xi_4}b^{\xi_2\xi_5}b^{\xi_3\xi_6}\left(\nu_{\xi_4,\xi_5}\nu_{\xi_6,h}[3]\right)\nu^{h,h} \\
&+\nu_{h\xi_1\xi_2}b^{\xi_1\xi_3}b^{\xi_2\xi_4}\nu_{\xi_4\xi_5\xi_6}b^{\xi_5\xi_7}b^{\xi_6\xi_8}\left(\nu_{\xi_7,\xi_8}\nu_{\xi_3,h}[3]\right)\nu^{h,h} \\
&+b^{\xi_1\xi_2}b^{\xi_3\xi_4}\nu_{h\xi_1,h\xi_3}\nu_{\xi_2,\xi_4}\nu^{h,h} \\
&+\frac{1}{4}\nu_{h\xi_1\xi_2}b^{\xi_1\xi_3}b^{\xi_2\xi_4}\nu_{h\xi_5\xi_6}b^{\xi_5\xi_7}b^{\xi_6\xi_8}\left(\nu_{\xi_3,\xi_4}\nu_{\xi_7,\xi_8}[3]\right)\nu^{h,h} \\
&+\nu^{h,\phi_1}\nu_{\xi_1,\phi_1,\phi_2}b^{\xi_1\xi_2}\nu^{\phi_2,h}\nu_{h\xi_3\xi_4}b^{\xi_3\xi_5}b^{\xi_4\xi_6}\left(\nu_{\xi_5,\xi_6}\nu_{\xi_2,h}[3]\right), \qquad (4.160) \\
D_1 \;=\;& \frac{1}{6}\left(\nu^{h,h}\right)^2\nu_{h,h,h,h} \\
&-\frac{1}{9}\left(\nu^{h,h}\right)^3\left(\nu_{h,h,h}\right)^2 \\
&+\nu^{h,\phi_1}\nu_{\xi_1,\phi_1,\phi_2}b^{\xi_1\xi_2}\nu^{\phi_2,h}\nu_{\xi_2,h,h} \\
&+\frac{4}{3}\nu^{h,\phi_1}\nu_{\xi_1,\phi_1,\phi_2}b^{\xi_1\xi_2}\nu^{\phi_2,h}\nu_{h,h,h}\nu_{\xi_2,h}\nu^{h,h} \\
&-2b^{\xi_1\xi_2}\nu_{h\xi_1,h,h}\nu_{\xi_2,h}\left(\nu^{h,h}\right)^2 \\
&+\frac{1}{3}b^{\xi_1\xi_3}b^{\xi_2\xi_4}\nu_{h\xi_1\xi_2}\nu_{\xi_3,\xi_4}\nu_{h,h,h}\left(\nu^{h,h}\right)^2
\end{aligned}$$

$$
\begin{aligned}
& -\frac{2}{3}b^{\xi_1\xi_3}b^{\xi_2\xi_4}\nu_{h\xi_1\xi_2}\nu_{\xi_3,h}\nu_{\xi_4,h}\nu_{h,h,h}\left(\nu^{h,h}\right)^3 \\
& +\frac{1}{2}b^{\xi_1\xi_3}b^{\xi_2\xi_4}\nu_{h\xi_1\xi_2}\nu_{\xi_3,h,h}\nu_{\xi_4,h}\left(\nu^{h,h}\right)^2 \\
& +\nu^{h,\phi_1}\nu_{\xi_1,\phi_1,\phi_2}b^{\xi_1\xi_2}\nu_{\xi_2\xi_3\xi_4}b^{\xi_3\xi_5}b^{\xi_4\xi_6}\nu^{\phi_2,h}\nu_{\xi_5,h}\nu_{\xi_6,h} \\
& +\nu^{h,\phi_1}\left(\nu_{\phi_1,\phi_2,\xi_1\xi_2}+2\nu_{\phi_1,\xi_1,\phi_2\xi_2}+\nu_{\phi_1,\phi_2,\xi_1,\xi_2}\right)b^{\xi_1\xi_3}b^{\xi_2\xi_4}\nu^{\phi_2,h}\nu_{\xi_3,h}\nu_{\xi_4,h} \\
& +2\nu^{h,\phi_1}\nu_{\xi_1,\phi_1,\phi_2}b^{\xi_1\xi_2}\nu^{\phi_2,\phi_3}\nu_{\xi_3,\phi_3,\phi_4}b^{\xi_3\xi_4}\nu^{\phi_4,h}\nu_{\xi_3,h}\nu_{\xi_4,h} \\
& -\frac{2}{3}\nu_{h\xi_1\xi_2\xi_3}b^{\xi_1\xi_4}b^{\xi_2\xi_5}b^{\xi_3\xi_6}\nu_{\xi_4,h}\nu_{\xi_5,h}\nu_{\xi_6,h}\left(\nu^{h,h}\right)^2 \\
& +2\nu_{h\xi_1\xi_2}b^{\xi_1\xi_3}b^{\xi_2\xi_4}\nu_{\xi_4\xi_5\xi_6}b^{\xi_5\xi_7}b^{\xi_6\xi_8}\nu_{\xi_3,h}\nu_{\xi_7,h}\nu_{\xi_8,h}\left(\nu^{h,h}\right)^2 \\
& +2b^{\xi_1\xi_2}b^{\xi_3\xi_4}\nu_{h\xi_1,h\xi_3}\nu_{\xi_2,h}\nu_{\xi_4,h}\left(\nu^{h,h}\right)^2 \\
& +\frac{1}{2}\nu_{h\xi_1\xi_2}b^{\xi_1\xi_3}b^{\xi_2,\xi_4}\nu_{h\xi_5\xi_6}b^{\xi_5\xi_7}b^{\xi_6\xi_8}\left(\nu_{\xi_3,\xi_4}\nu_{\xi_7,h}\nu_{\xi_8,h}[6]\right)\left(\nu^{h,h}\right)^2 \\
& -\nu_{h\xi_1\xi_2}b^{\xi_1\xi_3}b^{\xi_2\xi_4}\nu_{h\xi_5\xi_6}b^{\xi_5\xi_7}b^{\xi_6\xi_8}\nu_{\xi_3,h}\nu_{\xi_4,h}\nu_{\xi_7,h}\nu_{\xi_8,h}\left(\nu^{h,h}\right)^3 \\
& +2\nu^{h,\phi_1}\nu_{\xi_1,\phi_1,\phi_2}b^{\xi_1\xi_2}\nu^{\phi_2,h}\nu_{h\xi_3\xi_4}b^{\xi_3\xi_5}b^{\xi_4\xi_6}\nu_{\xi_5,h}\nu_{\xi_6,h}\nu_{\xi_2,h}\nu^{h,h} \\
& +\frac{1}{2}\nu^{h,\phi_1}\nu_{\xi_1,\phi_1,\phi_2}b^{\xi_1\xi_2}\nu^{\phi_2,h}\nu^{h,\phi_3}\nu_{\xi_3,\phi_3,\phi_4}b^{\xi_3\xi_4}\nu^{\phi_4,h}\nu_{\xi_2,\xi_4}\left(\nu_{h,h}\right)^2 \\
& +\nu^{h,\phi_1}\nu_{\xi_1,\phi_1,\phi_2}b^{\xi_1\xi_2}\nu^{\phi_2,h}\nu^{h,\phi_3}\nu_{\xi_3,\phi_3,\phi_4}b^{\xi_3\xi_4}\nu^{\phi_4,h}\nu_{\xi_2,h}\nu_{\xi_4,h}\nu_{h,h} \\
& +\nu^{h,\phi_1}\nu_{\xi_1,\phi_1,\phi_2}b^{\xi_1\xi_2}\nu^{\phi_2,h}b^{\xi_3\xi_5}b^{\xi_4\xi_6}\nu_{h\xi_3\xi_4}\left(\nu_{\xi_5,\xi_6}\nu_{\xi_2,h}[3]\right) \qquad (4.161)
\end{aligned}
$$

$$
\begin{aligned}
D_2 \;=\; & \frac{1}{9}\left(\nu^{h,h}\right)^3\left(\nu_{h,h,h}\right)^2 \\
& +\frac{4}{3}\nu^{h,\phi_1}\nu_{\xi_1,\phi_1,\phi_2}b^{\xi_1\xi_2}\nu^{\phi_2,h}\nu_{h,h,h}\nu_{\xi_2,h}\nu^{h,h} \\
& +\frac{2}{3}b^{\xi_1\xi_3}b^{\xi_2\xi_4}\nu_{h\xi_1\xi_2}\nu_{\xi_3,h}\nu_{\xi_4,h}\nu_{h,h,h}\left(\nu^{h,h}\right)^3 \\
& +\nu_{h\xi_1\xi_2}b^{\xi_1\xi_3}b^{\xi_2\xi_4}\nu_{h\xi_5\xi_6}b^{\xi_5\xi_7}b^{\xi_6\xi_8}\nu_{\xi_3,h}\nu_{\xi_4,h}\nu_{\xi_7,h}\nu_{\xi_8,h}\left(\nu^{h,h}\right)^3 \\
& +\nu^{h,\phi_1}\nu_{\xi_1,\phi_1,\phi_2}b^{\xi_1\xi_2}\nu^{\phi_2,h}\nu^{h,\phi_3}\nu_{\xi_3,\phi_3,\phi_4}b^{\xi_3\xi_4}\nu^{\phi_4,h}\nu_{\xi_2,h}\nu_{\xi_4,h}\nu_{h,h} \\
& +2\nu^{h,\phi_1}\nu_{\xi_1,\phi_1,\phi_2}b^{\xi_1\xi_2}\nu^{\phi_2,h}b^{\xi_3\xi_5}b^{\xi_4\xi_6}\nu_{h\xi_3\xi_4}\nu_{\xi_2,h}\nu_{\xi_5,h}\nu_{\xi_6,h}\nu^{h,h}. \qquad (4.162)
\end{aligned}
$$

By comparing (4.157) and (4.159) it can be seen that the coefficients of $\frac{1}{k!}t^k$ on the left- and right-hand sides of (4.143) are equal for all $k \geq 2$ if and only if $d = 2$ and

$$
\begin{aligned}
a_2 &= \frac{1}{4}D_2 && (4.163) \\
a_1 &= \frac{1}{2}D_1 - 2D_2 && (4.164) \\
a_0 &= D_0 - \frac{3}{2}D_1 - \frac{9}{4}D_2. && (4.165)
\end{aligned}
$$

Calculating similar expressions for the coefficients of t and the constant terms on the left- and right-hand sides of (4.143) shows that these are also equal if and only if $d = 2$ and a_0, a_1 and a_2 are as given by (4.165), (4.164) and (4.163) above.

Condition (4.143) holds if we take the expressions (4.165), (4.164) and (4.163) for a_0, a_1 and a_2 in equation (3.2) for the modified score test statistic. Thus we have a generalized Bartlett adjustment for the score test statistic under a mis-specified model, where the interest parameter is one-dimensional .

4.3 Extension of Generalized Bartlett Adjustment to dim $\psi > 1$ for Score and Likelihood Ratio Statistics

The derivations of the modified likelihood ratio and score test statistics under a mis-specified model in this chapter are algebraically complex, and consequently have only been carried out in the case where the interest parameter (ψ in the θ-parameterization or τ in the ϕ-parameterization) is one-dimensional. For the score test, there is no reason why the results should not be extended to the case where the interest parameter has dimension larger than 1. Cordeiro & Ferrari (1991) derive a formula for a generalized Bartlett adjustment for any statistic S which has distribution function

$$F_s(x) = G_r(x) + \sum_{i=0}^{k} a_i G_{r+2i}(x) + O\left(n^{-2}\right), \tag{4.166}$$

where $G_r(x)$ is the distribution function of a χ^2_r distribution and $\sum a_i = 0$. They also note the results of Chandra (1985), who proves that any statistic which, under the null hypothesis, has an asymptotic central chi-squared distribution with r degrees of freedom has a distribution function of the form (4.166).

Given the complexity of the expressions for the generalized Bartlett adjustment of the score statistic under a mis-specified model, the derivation of the adjustment for an arbitrary-dimension interest parameter would ideally be done using an approach which removes or minimizes the need for a particular parameterization. This might well have the additional benefit of revealing any underlying geometric meaning in the expressions.

It is unclear, however, whether equivalent expressions for the generalized Bartlett adjustment of the mis-specified likelihood ratio statistic with interest parameter of arbitrary dimension can be found. As stated earlier (see (4.97) and (4.98)), the asymptotic distribution of the likelihood ratio statistic under a mis-specified model is a weighted sum of χ_1^2 variates. When dim $\psi = 1$, we have, asymptotically as $n \to \infty$,

$$w \sim \mu U, \tag{4.167}$$

where U has a χ_1^2 distribution, and we may find a generalized Bartlett adjustment for $\mu^{-1}w$. No such approach is possible when dim $\psi > 1$.

In Section 3.2 we considered the weighted sum of two independent, asymptotically chi-squared statistics for which linear Bartlett adjustments exist. It was proved that generalized Bartlett adjustment of this weighted sum was possible only under rather restrictive conditions. It should be possible to generalize this argument to show that generalized Bartlett adjustment of the likelihood ratio statistic under a mis-specified model is possible only under similar conditions. Specifically, since the asymptotic distribution of the likelihood ratio statistic under a mis-specified model is a weighted sum of independent χ_1^2 variates, with the weights being the eigenvalues of the matrix given by (4.98), it seems likely from the example in Section 3.2 that the only non-trivial situation in which generalized Bartlett adjustment will be possible is that where all of the weights are equal. Kent (1982) discusses conditions for exponential families such that all the weights equal 1 and the likelihood ratio test statistic is robust.

4.4 Applications of Modified Score and Likelihood Ratio Test Statistics

In order to perform generalized Bartlett adjustment in practice on the score and likelihood ratio statistics under a mis-specified model it is necessary to estimate the joint log-likelihood cumulants $\kappa_{i,j}$, $\kappa_{i,jk}$, $\kappa_{ij,kl}$, etc. on which they depend. For the score test statistic, this does not present a major problem. The modified score test statistic is

$$S_R' = \left(1 + \frac{1}{n}\sum_{r=0}^{2} a_r S_R^r\right) S_R, \tag{4.168}$$

where the coefficients a_0, a_1 and a_2, given by (4.165), (4.164) and (4.163) respectively, depend on the log-likelihood cumulants. The a_r need only be estimated with error of order $O\left(n^{-\frac{1}{2}}\right)$, for if we have estimates $\hat{a}_r$ such that

$$a_r = \hat{a}_r + O\left(n^{-\frac{1}{2}}\right), \quad r = 0, 1, 2, \tag{4.169}$$

then

$$S_R' = \left(1 + \frac{1}{n}\sum_{r=0}^{2} \hat{a}_r S_R^r\right) S_R + O\left(n^{-\frac{3}{2}}\right). \tag{4.170}$$

which has a χ^2 distribution with error of order $O\left(n^{-\frac{3}{2}}\right)$ as desired. Thus derivation of the modified score statistic may be possible in practical situations, provided that the rather complex expressions for the coefficients a_r can be estimated for the required model. This should be possible with the help of computer algebra.

Unfortunately, the situation with the likelihood ratio statistic is not so simple. While it is sufficient to estimate the coefficients a_r to order $O(1)$ only, the generalized Bartlett adjustment when dim $\psi = 1$ is applied not to the mis-specified likelihood ratio statistic w, but to $\mu^{-1}w$ (see (4.167) above). This multiplier, μ^{-1}, must be estimated with error of order $O\left(n^{-\frac{3}{2}}\right)$ if the adjustment is to be applied. Estimation to this level of accuracy is likely to be difficult, if not impossible, in many practical situations.

References

[1] BARNDORFF-NIELSEN, O.E. (1986) *Strings, tensorial combinants, and Bartlett adjustments.* Proc. R. Soc. Lond. A, **406**, 127-37.

[2] BARNDORFF-NIELSEN, O.E. (1987) *Differential geometry and statistics: some mathematical aspects.* Indian J. Math., **29**, 335-50.

[3] BARNDORFF-NIELSEN, O.E., BLÆSILD, P., PACE, L. & SALVAN, A. (1991) *Formulas for Asymptotic Statistical Calculations.* Research Report No. 207, Dept. of Theoretical Statistics, University of Aarhus.

[4] BARNDORFF-NIELSEN, O.E. & COX, D.R. (1984) *Bartlett adjustments to the likelihood ratio statistic and the distribution of the maximum likelihood estimator.* J. R. Statist. Soc. B, **46**, 483-95.

[5] BARNDORFF-NIELSEN, O.E., JUPP, P.E. & KENDALL, W.S. (1994) *Stochastic calculus, statistical asymptotics, Taylor strings and phyla.* Annales de la Faculté des Sciences de Toulouse, **III**, 5-62.

[6] CHANDRA, T.K. (1985) *Asymptotic expansions of perturbed chi-square variables.* Sankhyā, **47**, 100-10.

[7] CORDEIRO, G.M. & FERRARI, S.L. DE P. (1991) *A modified score test statistic having chi-squared distribution to order n^{-1}.* Biometrika, **78**, 573-82.

[8] COX, D.R. & HINKLEY, D.V. (1974) *Theoretical Statistics.* Chapman and Hall, London.

[9] ERIKSEN, P.S. (1987) *Proportionality of Covariance Matrices.* Ann. Statist., **15**, 732-48.

[10] FRASER, D.A.S. (1965) *On information in statistics.* Ann. Math. Statist., **36**, 890-6.

[11] FRYDENBERG, M. & JENSEN, J.L. (1989) *Is the 'improved likelihood ratio statistic' really improved in the discrete case ?* Biometrika, **76**, 655-61.

[12] HARRIS, P. (1985) *An asymptotic expansion for the null distribution of the efficient score statistic.* Biometrika, **72**, 653-9.

[13] HOLLAS, E.A. (1991) *Performance of Bartlett adjustment for certain likelihood ratio tests.* Commun. Statist. -Simula., **20**, 449-62.

[14] KENT, J.T. (1982) *Robust properties of likelihood ratio tests.* Biometrika, **69**, 19-27.

[15] LAWLEY, D.N. (1956) *A general method for approximating to the distribution of the likelihood ratio criteria.* Biometrika, **43**, 295-303.

[16] MCCULLAGH, P. (1987) *Tensor Methods in Statistics.* Chapman and Hall, London.

[17] MØLLER, J. (1986) *Bartlett adjustments for structured covariances.* Scand. J. Statist., **13**, 1-15.

[18] MURRAY, M.K. & RICE, J.W. (1993) *Differential Geometry and Statistics.* Chapman and Hall, London.

[19] YFANTIS, E.A. & BORGMAN, L.E. (1982) *An Extension of the Von Mises distribution.* Commun. Statist. -Theor. Meth., **11**, 1695-706.